

Content Analysis

OLE R. HOLSTI, *University of British Columbia*

With the collaboration of Joanne K. Loomba and Robert C. North

The communication process is an intrinsic part of all social interaction from the interpersonal to the international level. Groups, institutions, and organizations, from the family to the nation, exist by virtue of communication and cease to exist once it becomes totally disrupted. It is thus axiomatic that the study of the processes and content of communication is basic to all social sciences, "for most social facts worth studying are embedded in a process of communication" (Pool, 1954, p. 352).

The interest of the social psychologist in the content of communication has been described by Cartwright (1953, p. 422):

When one stops to think of it, it is really surprising how much of the subject matter of social psychology is in the form of verbal behavior. The formation and transmission of group standards, values, attitudes, and skills are accomplished largely by means of verbal communication. Education in the schools, in the home, in business, in the neighborhood, and through the mass media is brought about by the transmission of information and by the exercise of controls which are largely mediated through written or spoken words. If one is concerned with problems of social organization, the situation is similar. Supervision, management, coordination, and the exertion of influence are principally matters of verbal interaction. Social and political conflicts, although often stemming from divergent economic interests and power, cannot be fully understood without studying the words employed in the interaction of conflicting groups, and the process of mediation consists largely in talking things out. The work of the world, and its entertainment, too, is in no small measure mediated by verbal and other symbolic behavior.

Useful comments and suggestions on an earlier draft of this chapter were offered by Richard R. Fagen, Bo Ohlström, Kenneth C. Prewitt, Philip J. Stone, Lois Swirsky, and Robert G. Weinland. Their assistance is gratefully acknowledged, as is that of Mrs. Arlee Ellis and Mrs. Violet Lofgren, whose contributions are too numerous to list exhaustively.

These comments are equally appropriate for other social sciences. As a consequence, the study of communication content has been approached from a variety of different starting points and undertaken with the tools and conceptual frameworks of several disciplines.

Content analysis is a multipurpose research method developed specifically for investigating a broad spectrum of problems in which the content of communication serves as the basis of inference. In the introductory section of this chapter we discuss specific definitions of content analysis, types of content-analysis research designs, and major trends in the development of content analysis. Within the framework developed in the introduction, the second section surveys research in which content analysis has been employed. This is followed by an examination of some problems of research design: units and categories of analysis, and types of measurement. The fourth section is concerned with the problems of sampling, reliability, and validity. The chapter concludes with a discussion of content analysis by means of electronic computers, a recent development with important implications for many types of research problems.

DEFINITION OF CONTENT ANALYSIS

Nearly all research in the social sciences and humanities depends in one way or another on careful reading of written materials. Given the ubiquity of this process in research, what characteristics distinguish content analysis from any careful reading of documents? Definitions of content analysis have tended to change over time with developments in technique and with the application of the tool itself to new problems and types of materials. Among the definitions which have been proposed are the following:

Content analysis is the statistical semantics of political discourse. (Kaplan, 1943, p. 230)

"Content analysis" may be defined as referring to any technique (a) for the classification of the sign-vehicles, (b) which relies solely upon the judgments (which theoretically, may range from perceptual discriminations to sheer guesses) of an analyst or group of analysts as to which sign-vehicles fall into which categories, (c) on the basis of explicitly formulated rules, (d) provided that the analyst's judgments are regarded as the reports of a scientific observer. (Janis, 1949, p. 55)

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

We propose to use the terms "content analysis" and "coding" interchangeably to refer to the objective, systematic, and quantitative description of any symbolic behavior. (Cartwright, 1953, p. 424)

The term "content analysis" is used here to mean the scientific analysis of communications messages. . . . The method is broadly speaking the "scientific method," and while being catholic in nature, it requires that the analysis be rigorous and systematic. (Barcus, 1959, p. 8)

Content analysis refers to any procedure for assessing the relative extent to which specified references, attitudes, or themes permeate a given message or document. (Stone, 1964)

This selective sampling of definitions indicates that, along with a persisting consensus about some characteristics, there has been a marked tendency to broaden the boundaries of content analysis by means of less restrictive definitions.

Among the characteristics of content analysis on which there is wide agreement are those of *objectivity*, *system*, and *generality*. To have *objectivity*, the analysis must be carried out on the basis of explicitly formulated rules which will enable two or more persons to obtain the same results from the same documents. In a *systematic* analysis the inclusion and exclusion of content or categories is done according to consistently applied criteria of selection; this requirement eliminates analyses in which only materials supporting the investigator's hypotheses are examined. By *generality* we mean that the findings must have theoretical relevance; purely descriptive information about content, unrelated to other attributes of content or to the characteristics of the sender or recipient of the message, is of little scientific value. These three requirements are not unique to content analysis, but are necessary conditions for all scientific inquiry. They serve to indicate that, in general terms, content analysis can be regarded as the application of the principles of scientific research to the analysis of communication content.

Alongside general consensus on the defining characteristics of content analysis—objectivity, system, and generality—considerable debate on two other requirements has been generated in recent literature. First, must content analysis be *quantitative*? And second, must it be limited to the *manifest* content, or may it be used also to probe for the more latent aspects of communication?

Quantification has usually been accepted as one of the most important characteristics of content analysis. According to one standard source, "There is clearly no reason for content analysis unless the question one wants answered is quantitative" (Lasswell, Lerner, and Pool, 1952, p. 45). However, there is considerable disagreement about the meaning of the term "quantitative." Some definitions equate it with *numerical*: "Content analysis aims at a classification of content in more precise, *numerical terms* than is provided by impressionistic 'more or less' judgments of 'either-or'" (Kaplan and Goldsen, 1949, p. 83). Other definitions are less restrictive and include studies in which findings are reported in such terms as "more," "less," or "increasing" (Berelson, 1952, p. 17). However, qualitiveness and quantitiveness are not dichotomous attributes, but fall along a continuum (Lazarsfeld and Barton, 1951). To state that attribute *A* became more important in *X*'s messages with the passing of time is a qualitative assertion, but it is not without a quantitative aspect; nor is such a statement as "In document *Y*, *A* asserts *X*" purely qualitative. Data reported in this form imply ordinal and nominal scaling. Moreover, statistical methods can be used with such data.

Definitions of content analysis requiring that inferences be derived strictly from counts of *frequency* (for example, Leites and Pool, 1942, pp. 1-2; Janis, 1943, p. 429) place a number of standard methods on the borderline of acceptability. A pioneering application of content analysis, the RADIR studies, combined frequency and nonfrequency techniques; each editorial in the sample

taken from a series of "prestige newspapers" during a sixty-year period was coded according to the appearance or nonappearance of certain key symbols (Lasswell, Lerner, and Pool, 1952). Thus a strict frequency count, including multiple use of any symbol within an editorial, was not employed. A similarly "mixed" characteristic is found in the technique of *contingency analysis*, in which the coding of material depends on the absence or presence of the attribute within some content unit, rather than on the frequency of its presence (Osgood, 1959, p. 63).

The case for content analysis based on exact counts of frequency is a powerful one. Foremost among the arguments is the degree of precision with which one's conclusions may be stated. Descriptions such as "45 percent" or "27 times out of a possible 30" convey information more precisely than statements such as "less than half" or "almost always." In response to the question "Why quantify?" Lasswell, Lerner, and Pool (1952, pp. 31-32) have pointed to a number of other questions which often remain unanswered in qualitative symbol studies:

Can we assume that a scholar read his sources with the same degree of care throughout his research? Did he allow his eye to travel over the thousands of pages of parliamentary debates, newspapers, magazines and other source lists in his bibliography or notes? Or did he use a sampling system scanning some pages superficially, though concentrating upon certain periods? Was the sampling system for the *Frankfurter Zeitung*, if one was employed, comparable with the one for the *Manchester Guardian*? Were the leaflets chosen simply because they were conveniently available to the scholar, or were they genuinely representative of the most widely circulated propaganda leaflets?

Finally, statistical methods, which permit a more accurate description of the degree of covariation of two or more attributes, require some degree of quantification. They also permit a more precise answer to a recurring question raised by case studies: with what degree of confidence can one generalize from the results obtained in the sample under study? But the use of statistics is not dependent solely on frequency counts or any other single system of enumeration.

Despite the widely recognized advantages of quantitative methods, the tendency to equate content analysis with tabulation of frequencies has been criticized on a number of grounds. The most general of these is the charge that such a restriction leads to bias in the selection of problems to be investigated, undue emphasis being placed on precision at the cost of problem significance (Barcus, 1959; Smythe, 1952).

Related to this general criticism is the view that one can draw more meaningful inferences by nonquantitative methods (Kracauer, 1952). Qualitative content analysis, which has sometimes been defined as the drawing of inferences on the basis of appearance or nonappearance of attributes in messages, has been defended largely, though not solely, for its superior performance in problems of applied social science (George, 1959b, pp. 9-10). When content from propaganda sources is used to predict enemy behavior, the pressure of time, the inability to control variables, and the possibility that nonrecurring phenomena may provide major clues to policy often render exhaustive quantitative analyses uneco-

nical and difficult, if not impossible, to design and carry out. Citing instances in which qualitative analysts were able to draw more accurate inferences from studies of Nazi propaganda during World War II than could those using quantitative techniques, George concluded: "Qualitative analysis of a limited number of crucial communications may often yield better clues to the particular intentions of a particular speaker at one moment in time than more standardized techniques" (1959b, p. 7).

In line with this reasoning, proponents of qualitative techniques also question the assumption that, for purposes of inference, frequency of assertion is necessarily related to the importance of the assertion; these critics suggest that the single appearance or omission of an attribute in a document may be of more significance than the relative frequency of other characteristics (George, 1959b). An example of this point is found in a study of Chinese documents prior to China's active entry into the Korean war in October 1950. The change from the passive term *fan tui* to the word *k'ang yi*, previously used as an exhortation to action against Japan and against the Chinese Nationalists, provided the first clue that Chinese leaders had decided to intervene in the war (Whiting, 1960, p. 99). But even studies which emphasize the unique aspects of each document are not simply qualitative; rather than counting frequencies, the analysts have chosen to formulate nominal categories into which one of two scores is recorded—present or absent.

Finally, whether stated explicitly or not, even the most rigorously quantitative study uses qualitative techniques at some stage in the research, either as a preliminary step in determining which variables are most likely to prove useful when measured, or as a final check on the face validity of the findings. Pool (1959, p. 192) summarizes this point: "It should not be assumed that qualitative methods are insightful, and quantitative ones merely mechanical methods for checking hypotheses. The relationship is a circular one; each provides new insights on which the other can feed."

It is worth noting in conclusion that for most scientific research the advantages to be gained by some type of quantification continue to be important. Nevertheless, for purposes of definition there are few compelling reasons for excluding studies which fail to conform to any single system of enumeration from the proper jurisdiction of content analysis.

A second major source of disagreement among those defining content analysis is whether it must be limited to manifest content, that is, the surface meaning of the content. Or may content analysis be used to analyze the deeper layers of meaning embedded in the content? The manifest-latent controversy can be considered at two levels. The requirement of objectivity stipulates that only those symbols and combinations of symbols actually appearing in the message be recorded. In other words, the *coding* process cannot be one of "reading between the lines." In this sense, content analysis is limited to manifest attributes of text.

The second aspect of the manifest-latent issue concerns the *interpretation* of results. This debate is essentially one concerning the dimensions of communication which may properly be analyzed (Morris, 1946). Earlier definitions tended

to limit content analysis to questions of semantics, the relationship of signs to referents, and to questions of syntactics, the relationship of signs to signs (Kaplan, 1943; Janis, 1949; Berelson, 1952). The restriction against analysis of the pragmatic dimension of language, the relationship of signs to those that produce or receive them, was usually based on the difficulty of drawing valid inferences about the causes or effects of communication directly from content data.

As has been the case in the quantitative-qualitative debate, the recent trend has been in the direction of a broader definition (Cartwright, 1953, p. 424; Barcus, 1959, p. 19; Dunphy, 1964). Nearly the entire volume of papers from the Work Conference on Content Analysis of the Social Science Research Council was addressed to using messages for purposes of answering questions about the causes or effects of communication (Pool, 1959). This trend toward a broader view of content analysis is evident in Osgood's definition (1959, p. 36):

... we define content analysis as a procedure whereby one makes inferences about sources and receivers from evidence in the messages they exchange. . . . [W]hen the interest of the content analyst lies in making inferences about the source of a message, he must rely upon encoding dependencies, that is, the dependencies of message events upon psychological processes in speakers and writers. When his interest lies in making inferences about the effects of a message upon its receivers, on the other hand, he relies upon decoding dependencies, that is, the dependencies of events in listeners and readers (their meanings, emotions, attitudes, and the like) upon the content and structure of messages.

In this view, content analysis is considered an integral part of the field of psycholinguistics (Osgood and Sebeok, 1954; Jaffe, 1966).

Actually, the differences between the broader and more restrictive views are not so great as suggested at first glance. Both Kaplan (1943, p. 223) and Janis (1943, p. 437) exclude pragmatic content analysis because inferences as to the causes or effects of content can rarely be validated solely by analysis of the messages themselves. On the other hand, proponents of a broader definition usually assume that content data will be used, directly or indirectly, in conjunction with independent behavioral indices.

For present purposes, a broad definition of the field will be adopted: *Content analysis is any technique for making inferences by systematically and objectively identifying specified characteristics of messages.* While any definition adopted is open to criticism, one which errs on the side of inclusiveness permits a broader consideration of the literature, and thus has an advantage over one which is overly restrictive.

WHEN TO USE CONTENT ANALYSIS

When should one use content analysis? The wide range of possibilities is suggested by reference to a few of the diverse problems which have been subjected to content-analysis techniques in recent years:

1. *What differences among cultures are reflected in the songs and literature of various nations (Sebald, 1962; Lewin, 1947)?*

2. To what extent have the political symbols of the New Deal been adopted by American conservatives (Prothro, 1956)?
3. What differences characterize the language behavior of schizophrenic and normal persons (Mann, 1944; Fairbanks, 1944)?
4. Is editorial support of a political candidate also reflected in biased news sections (Kobre, 1953; Klein and Maccoby, 1954)?
5. What has been the Soviet reaction to Voice of America broadcasts (Inkeles, 1952)?
6. Is the Riesman hypothesis about the increasing other-directedness of American society supported by changing content of consumer goods advertising (Dornbusch and Hickman, 1959)?
7. How is sentence length related to the comprehensibility of literature (Coleman, 1962)?
8. How are expressions of the "need for achievement" related to stages in the development of a civilization (McClelland, 1961; deCharms and Moeller, 1962)?
9. Who was the author of the Federalist Papers, Nos. 49-58, 62, and 63 (Mosteller and Wallace, 1964)?

Content analysis is not relevant to all documentary research, however. It can rarely be used to determine the truth of an assertion, or to evaluate the esthetic qualities of poetry. Moreover, if the social scientist uses documents to settle limited issues of fact, such as to determine which newspapers supported Kennedy and Nixon in 1960, methods other than content analysis could be used more efficiently. But the investigator's questions and the content of documents are rarely coterminous (Dibble, 1963, p. 216). One approach to documentary research is exemplified by a recent manual, which suggests dependence on "a sort of sixth sense that will alert you to tell-tale signs" (quoted in Dibble, 1963, p. 204). Content analysis may be considered as a supplement to, not a substitute for, subjective examination of documents. On the other hand, the limitation of depending solely on ordinary reading of documents for purposes of scientific research are illustrated in a study of Richard Wright's autobiography, *Black Boy* (White, 1947). Although the investigator was a trained psychologist, his preliminary appraisal of the book failed to uncover a number of major themes. Systematic content analysis, however, revealed Wright's emphasis on personal safety (18 percent of all value judgments), failure to identify with other Negroes, and lack of interest in social goals. In general, then, content analysis is useful whenever the problem requires precise and replicable methods for analyzing those aspects of symbolic behavior which may escape casual scrutiny.

More specifically, content analysis is likely to prove especially useful for at least three general classes of research problems. Content analysis may be helpful when there are technical advantages because the volume of material to be examined is such that the investigator must either (1) confine his study to some sample

of the total universe of communication, (2) use a team of assistants, each with his own subjective predispositions, or (3) both. For example, an analysis of key symbols appearing in "elite" newspapers over a sixty-year period (Lasswell, Lerner, and Pool, 1952), could not be undertaken on a nonsystematic basis; nor could every issue of each newspaper for the entire period be examined. Such a study clearly required the use of both sampling and many research assistants. Unless rigorous and explicitly formulated rules for systematic sampling and reliable coding were used, inferences drawn from these data would be of questionable value. Content analysis is a technique for meeting these requirements.

Content analysis may also prove useful when data accessibility is a problem and the investigator's data are limited to the messages produced by individuals. Restrictions of time or space often do not permit the social scientist to gain direct access to his subject. In such cases the subject must be studied "at a distance," with the consequence that other social-science research techniques (interview, questionnaire, observation) are not applicable. If the subject is no longer alive, he can be studied only through the record of his activities, through what his contemporaries set down about him, or through whatever writings he has left. In some instances the third category constitutes the most revealing, and occasionally the only surviving, source. Thus content analysis can serve as a "last resort" approach to social research when more direct techniques of analysis are ruled out by circumstances. Identification of unknown authors (Yule, 1944; Paisley, 1964), inferences from enemy propaganda (George, 1959a; Lasswell, Leites *et al.*, 1949; Whiting, 1960), or analysis of decision makers' attitudes during international crisis (Holsti, Brody, and North, 1965) illustrate this use of content analysis.

Finally, some form of content analysis is often necessary when, given certain theoretical components of the data themselves, the subject's own language is crucial to the research problem. To analyze psychiatric interviews, projective tests, or open-ended questionnaires, the social scientist often requires information of a subtlety or complexity which renders casual scrutiny inadequate, even if undertaken by a skilled and sensitive reader. One would have little confidence in rough estimates of the degree of "need achievement" in Indian folklore (McClelland and Friedman, 1952), or of the type/token ratio—the number of different words in text of a given length—in the language of schizophrenics (Fairbanks, 1944; Page, 1953), regardless of the investigator's skill or training.

CONTENT-ANALYSIS RESEARCH DESIGNS

All communication is composed of six basic elements: a *source* or sender; an *encoding process* which results in a *message*; a *channel* of transmission; a *detector* or recipient of the message; and a *decoding process*. Although content analysis is always performed on the message, it may be used to answer questions about each of the other elements of communication. The classical formulation of these questions is, "Who says what, to whom, how, and with what effect?" (Lasswell, Lerner, and Pool, 1952, p. 12), to which one more questions—"Why?"—may be added.

Each of these questions may be subsumed under research designed for three different purposes. The investigator may analyze messages to test hypotheses and make inferences about (1) characteristics of the text, (2) causes or antecedents of the message, or (3) effects of the communication. These three categories differ with respect to the questions asked of the data, the dimension of communication analyzed, and the research design. These differences are summarized in Table 1. Here it will be useful to consider problems and implications inherent in each of the three categories.

The first and most frequent application of content analysis has been for the purpose of *describing the attributes of the message*, without reference to either the intentions (encoding process) of the sender or the effect of the message on those to whom it is directed (decoding process). Much of the research has addressed itself to some variety of the "what" question, testing hypotheses about such matters as focus of attention, trends in communication, or cross-media differences. The investigator may also want to answer the question "to whom," as when testing hypotheses about the way in which the content of messages will differ according to audience (Berkman, 1963), or to answer questions about "how,"

for example, studies of style (Skinner, 1939) or of techniques of persuasion (Lee, 1952).

The type of research design will depend on the questions the investigator seeks to answer and on his data. In order to state meaningful conclusions, however, *all* content data must be compared to some other data. To determine that an editorial used the term "freedom" X number of times is a meaningless finding by itself. When content analysis is used to describe text, there are three basic types of comparison to be made.

The analyst may compare documents derived from *a single source*. One application of this method is the comparison of the messages over *time*. Such studies have been undertaken to analyze trends in the development of a discipline (Allport and Bruner, 1940; Tannenbaum and Greenberg, 1961b), and changes in the content of a rural newspaper (Taeuber, 1932) or Negro attitudes toward discrimination (Rosen, 1964). The investigator may also compare messages from a single source in differing *situations*, as in a study of international communication under conditions of low and high stress (Holsti, 1965b). Another, less widely used, approach using content data from a single source involves comparisons across *audience*. Comparisons of communication attributes across time, situation, or audience are intermessage analyses. Inferences may also be based on the *covariation of two or more variables*, such as self concepts and feeling toward others (Sheerer, 1949; Stock, 1949), within a single document or set of documents.

Hypotheses may be tested by comparing the messages of *two or more different sources*. Usually the purpose is to relate theoretically significant attributes of communication sources to differences in the messages they produce. For example, this design has been used to determine whether competitive newspapers provide significantly better news coverage than noncompetitive ones (Willoughby, 1955; Nixon and Jones, 1956), to examine linguistic differences in messages produced by "normal" persons and those on the point of committing suicide (Osgood and Walker, 1959), or to identify differences in propaganda appeals of two or more orators (Shneidman, 1963), political parties (Almond, 1954), or nations (Lasswell, 1927).

Finally, content data may be compared to some *standard* of adequacy or performance. Many studies have employed *a priori* standards defined, often implicitly, by the investigator's preferences. Even when *a priori* standards are made explicit, the problems of defining operationally such deviations as "bias" have rarely been dealt with in a satisfactory manner. An alternative to the deductive approach is to derive standards inductively from *content data*. A representative sample of messages produced by a class of communicators may provide norms against which the products of any single communicator may be compared. This technique has often been used in mass media research: the content of a given newspaper, magazine, or network is compared to the performance of the medium as a whole. A third type of standard against which content data may be compared is one defined by *noncontent indices*, such as aggregate data or expert opinion. A classic study of this type compared the incidence of minority group characters in popular magazine fiction with census data (Berelson and Salter, 1946).

TABLE 1
CONTENT-ANALYSIS RESEARCH DESIGNS

Purpose	Branch of semiotics	Types of comparisons	Questions	Research problem
To describe characteristics of communication	Semantics (sign/referent)	Messages source type A 1. Variable X across time 2. Variable X across situations 3. Variable X across audience 4. Variables X and Y within same universe of document	What?	To describe trends in communication content To relate known characteristics of sources to the messages they produce To audit communication content against standards
			How?	To analyze techniques of persuasion To analyze style
	Syntactics (sign/sign)	Messages source type A/ messages source type B	To whom?	To relate known characteristics of the audience to messages produced for them To describe patterns of communication
To make inferences as to the antecedents of communication (the encoding process)	Pragmatics (sender/sign)	Messages/nonsymbolic behavioral data 1. Direct 2. Indirect	Why?	To secure political and military intelligence To analyze psychological traits of individuals To infer aspects of culture and cultural change To provide legal evidence
			Who?	To answer questions of disputed authorship
To make inferences as to the effects of communication (the decoding process)	Pragmatics (sign/receiver)	Sender messages/recipient messages Sender messages/recipient behavioral data	With what effect?	To measure readability To analyze the flow of information To assess responses to communication

Many research designs use two or more types of comparison together. For example, hypotheses about the relative potency of role and personal variables in decision making were tested by analyzing the statements of Democratic and Republican Senators (comparisons across sources) during the Truman and Eisenhower Administrations (comparisons across situations) (Rosenau, 1968).

The second major classification of studies is that in which the text is analyzed in order to make *inferences about the causes or antecedents of the message*, and more specifically, about the sender. Thus content analysis is employed to discover "lawful relations *between* events in messages and processes transpiring in the individuals who produce . . . them" (Osgood, 1959, p. 36). Within the communication paradigm, messages are examined for the purpose of answering the questions "who" and "why." Who was the author of a given document? What are the meanings, associations, values, motives, or intentions of the communicator which can be inferred from his messages? Whereas the description of text can be classified under semantics or syntactics, this use of content analysis is a problem in pragmatics, the relationship of signs to those who produce them.

In order to draw valid inferences about sources from the messages they send, the content data must be compared, *directly* or *indirectly*, with independent behavioral indices. Because of possible differences in encoding habits—words may have different semantical meanings for different sources—inferences as to the antecedent causes of messages drawn solely from content data cannot be considered self-validating.

The relationship of symbolic behavior and its causes may be established by comparing content-analysis data *directly* with some other indices of behavior, as in a study of the 1914 crisis in which the degree of hostility in documents written by European decision makers was compared to the level of violence in their military actions (Holsti, North, and Brody, 1968).

Inferences based on an *indirect* relationship between symbolic and other forms of behavior are much more frequent in the content-analysis literature. The logic of such inferences can be stated as a syllogism: In a given situation, individuals whose behavior patterns are known to be *A*, *B*, and *C* produce messages with characteristics *r-s-t*, *u-v-w*, and *x-y-z*, respectively. If in similar circumstances a source produces messages with attributes *x-y-z*, the inference is that it was related to behavior pattern *C*. A classic example of this research design is the comparison between Nazi propaganda themes with the books, periodicals, and transoceanic cables of certain domestic organizations suspected of sedition, in which content analysis revealed significant similarities on a number of dimensions (Lasswell, 1949). The data obtained by content analysis were admitted by the court as legal evidence supporting the charge of sedition. In this case the likenesses in the messages of *two separate sources* served as the basis for inferences about the similarity of motives—support for the German war effort.

The research design may also focus on the relationship of events to symbols for a *single source*. The intelligence analyst may examine enemy documents for attributes that have in the past provided a clue to some aspect of the enemy's behavior. In general, the investigator bases his inferences on some demonstrated

relationship between events and symbols for the same or for comparable communicators. One weakness of many content-analysis studies is that such a relationship between symbolic behavior and other forms of behavior has not been established.

The third major classification of content-analysis studies is that in which *inferences are made about the effects of messages* (the decoding process) on the recipient. The question "with what effect" is a crucial aspect of the communication paradigm; it includes such a traditional concern of social psychology as attitude change. Nevertheless, relatively few studies have attempted to answer this question by content analysis. Probably the most systematic research on the decoding process has focused on research to measure the readability of text.

As in the other types of research design, the sender's message serves as the data. Two types of comparison may be used to measure the impact of the message. First, the investigator may determine the effects of *A's* messages to *B* by content-analyzing *B's* messages. For example, reviews of a book have been analyzed to investigate the responses it elicits (Lerner, 1950). An alternative approach to studying the effects of communication is to examine other aspects of the recipient's behavior. In studies of readability, for example, various independent tests have been used to measure the extent of the reader's comprehension.

TRENDS IN CONTENT-ANALYSIS RESEARCH

The history of content analysis as a research technique dates from the beginning of the twentieth century, although scattered studies going as far back as the 1740's have been cited (Dovring, 1954). During this time the method has gone through a number of phases. The description that follows relies in part on the findings presented in a study of content analysis from 1900 through 1958 (Barcus, 1959).

The most evident trend in content analysis is an almost geometric increase in the frequency of such research. During the first two decades of the century, an average of approximately 2.5 content-analysis studies appeared each year. During the next three decades, the annual average frequencies rose to 13.8, 22.8, and 43.3 respectively, and by 1950–1958, this figure had more than doubled again to 96.3. Or, to state this somewhat differently, there were more studies employing content analysis in the nine years after 1950 than during the previous half-century. Even a casual survey of the literature suggests that the trend toward increased content-analysis research has not abated.

A more revealing view of the history of content analysis can be presented when the studies are classified according to purpose, discipline or approach, and sources examined. Empirical studies account for an overwhelming proportion of the content-analysis literature. Barcus, in surveying 1719 titles, found that nearly four out of five could be so classified, the remainder being divided between methodological research (for example, studies of sampling or reliability in content analysis) (14.2%) and studies of form or style (6.9%). This three-fold classification underestimates the concern of content analysts with methodological problems; unquestionably many works listed as empirical studies of content have explicitly

come to grips with such issues. Nevertheless, concern with methodological issues has been notably more evident recently than in earlier content-analysis research. During the first two decades of the twentieth century, more than 98 percent of the studies dealt primarily with the content of text. By the 1950's, this figure had dropped to 75 percent, whereas one study in six was concerned with methodological problems. The recent development of various techniques of computer content analysis has served as an added impetus to the consideration of theoretical and methodological issues. This trend has not been without its critics, however. On the one hand, content analysts have been accused of being overly concerned with developing techniques, to the neglect of testing significant communication hypotheses (Stephenson, 1963). On the other hand, they have been charged with a lack of serious attention to theoretical and methodological defects (Krippendorff, 1966).

Another way to describe trends in content-analysis research is by approach or discipline. Three disciplines have accounted for approximately three-quarters of all empirical studies: sociological-anthropological (27.7%), general communications (25.9%), and political (21.5%). Using more specific categories, Barcus found that over three-fifths of all empirical content-analysis research has been focused in five mutually exclusive areas of inquiry, each of which accounts for at least 10 percent of the total: the study of social values, propaganda analyses, journalistic studies, media inventories, and psychological-psychoanalytic research.

Various phases have characterized the history of content analysis. Early investigations were confined largely to media inventories and journalistic studies, most of them devoted to studies of general American dailies. On the other hand, relatively little attention was paid to the foreign press. During the 1930's newspaper research continued to account for the largest number of studies. At the same time content analysis was increasingly being adapted for sociological, historical, and political research. The latter category included studies of propaganda—many of them stimulated by the seminal work of Harold Lasswell and his associates—foreshadowing a trend which gained added impetus during World War II. In the course of the 1940's, political research using propaganda materials accounted for nearly a quarter of all empirical content-analysis investigations. The dominating influence of political studies during the period of World War II is reflected in one of the definitions cited earlier, in which content analysis was characterized as “the statistical semantics of political discourse” (Kaplan, 1943, p. 230). Newspapers continue to be the source most frequently examined; but whereas such research accounted for over 60 percent of the studies prior to 1920, the figure had dropped to less than 30 percent by the 1950's. The development of new mass media (movies, radio, television) has also stimulated considerable content-analysis research.

Although journalistic and political studies continue to appear with increasing frequency, the most discernible trend has been toward wider application within a variety of disciplines. One of the milestone publications in the history of content analysis, appearing at the end of the 1950's, reveals something of the

diversity which has come to characterize the applications of the technique within such disciplines and spheres of specialized inquiry as folkloristics, biography, history, psychoanalysis, linguistics, propaganda, cognitive organization, and psychotherapy (Pool, 1959). These studies also clearly indicate a shift of interest away from descriptive studies toward research for testing hypotheses, especially regarding the antecedents and effects of communication. In recent years there have also been some successful attempts to use content analysis in areas beyond the traditional boundaries of communication by written or spoken word—for example, the analysis of pictorial communication (Gordon, 1952; Wayne, 1956; McClelland, 1958; Badri and Dennis, 1964), studies of bodily gesture and facial expression (Ekman, 1965), vocal tone (Starkweather, 1956), and even dreams (Hall and Van de Castle, 1966).

Finally, earlier uses of content analysis were largely confined to analyses of “natural” or “available” data—that is, data which exist without any active participation by the social scientist, such as newspapers, books, government documents, and personal documents.

Increasingly, however, content analysis has also been applied to messages generated in the course of the research process. This class of data may be of two somewhat different types. There is, in the first place, the research method generating only verbal data, which may then be subjected to content analysis. The psychiatric interview and various projective instruments, such as the Thematic Apperception Test, are representative examples. A second type of data is the by-product of other standard techniques of social research, which the investigator may then wish to content-analyze. Among the latter we may include responses to open-ended questions generated in survey research (Scheuch and Stone, 1964), written messages derived from a simulation study (Brody, 1963), or verbal communication produced during group interaction (Bales, 1950; Mills, 1964).

One inevitable result of this trend toward analysis of more diverse data has been a weakening of rigid boundaries between content analysis and other techniques of social research. The coding of open-ended questionnaires, for example, can often be considered to fall under the rubric of both content analysis and survey research.

In summary, the history of content analysis reveals a series of interrelated trends toward:

1. *Increased use of content analysis.*
2. *Heightened concern for theoretical and methodological issues.*
3. *Application to a broader spectrum of problems, especially those focusing on the antecedents and effects of communication.*
4. *Increased use for testing hypotheses, as opposed to purely descriptive research.*
5. *Greater diversity in the materials studied.*
6. *Use in conjunction with other techniques of social research.*
7. *Content analysis by means of computers.*

THE USES OF CONTENT ANALYSIS

Content analysis encompasses a widely varying set of techniques which have been employed to analyze symbolic materials ranging from propaganda tracts to Greek vases. Its scope and method can be described in greater detail through examination of problems to which it has been applied, either alone or in conjunction with other methods of social research.

One disclaimer must be made explicit at the outset: limitations of space restrict our review to a small sample of the total content-analysis literature. Only a few studies of each type can be described, and then only in summary fashion. Citations keyed to other studies listed in the reference section will be used to supplement the review, but even this bibliography can encompass only a fraction of the entire literature. Even though the scope of content analysis has been defined broadly, its application to some of the less frequently studied aspects of communication—handwriting, nonverbal communication, and content-free communication—will be excluded. The reader interested in these might turn to articles by Wolfson (1951), Ekman (1965), and Starkweather (1956).

This summary review is organized according to the categories depicted in Table 1. Research is initially divided according to major purpose: (1) to describe the characteristics of content; (2) to make inferences about the causes of content; and (3) to make inferences about the effect of content. Within this trichotomy studies are classified by type of question in the communication paradigm (what, how, to whom, why, who, with what effect), and are further subdivided according to the subgroups in the right-hand column of Table 1.

CHARACTERISTICS OF CONTENT: WHAT

Content analysis has been used most frequently for research problems in which the question can be answered *directly* from a description of the attributes of content. In such studies the investigator is in large part freed from problems of validity, except to the extent that validity is related to sampling and reliability; the content data serve as a direct answer to the research question, rather than as an indicator from which other characteristics are to be inferred.

To describe trends in communication content

Interest in measuring trends in national attitudes relevant to international politics has stimulated considerable analysis of the printed media. An early study applied scaling techniques developed by Thurstone to Chinese and Japanese materials during the Far Eastern crisis of 1930–1932 (Russell and Wright, 1933). The same techniques were used to examine trends in American attitudes toward Japan and China along a favorable-hostile continuum (Wright and Nelson, 1939).

After the outbreak of World War II, Lasswell (1941) suggested that a “world attention survey” be conducted by ongoing analyses of the world’s press. This idea subsequently led to an extensive survey of political symbols as one aspect of research on Revolution and Development of International Relations (RADIR).

The studies were designed to test hypotheses relating to a “world revolution” by identifying and mapping trends in the usage of those symbols expressing major goal values of modern politics. Editorials from one “prestige” newspaper representing each of five countries—the United States, Great Britain, France, Germany and the Soviet Union—were analyzed for the period 1890–1949 (Pool, 1952a). The choice of newspapers rather than other sources of political symbols was based on both theoretical and practical considerations (Lasswell, Lerner, and Pool 1952, p. 17):

In many countries the head of state makes few public statements, and these may be almost purely ceremonial. Party platforms often go unrevised year after year. So far as pressure groups are concerned, there are great differences in importance from state to state, and the task of gathering such fugitive material is a vast research project. The published proceedings of legislative bodies are often scanty, or altogether lacking. All states are modern enough to publish newspapers, however, so that comparable channel can be used for comparative purposes. Further, we can be relatively sure of who controls and who reads the news. In nearly all states some papers are understood to be leading organs of the party in office and of the government. Where the party system is competitive, the principal organ of parties and factions can be identified. Even with the coming of radio and television, the daily paper continues to occupy an important position in the media (and newspaper content is more accessible to study than are broadcasts).

Editorials appearing on the first and fifteenth day of each month were coded for the presence of 416 key symbols. These included 206 geographical terms, such as names of countries, international organizations, and minority groups, and 210 major ideological and doctrinal symbols—democracy, equality, proletariat, communism, nazism, nationalism, fatherland, and the like—relating to world politics in the first half of the twentieth century. Each time a symbol appeared it was scored as present, and furthermore, expressed attitudes toward the symbol—approval, disapproval, or neutrality—were recorded. Frequency counts were based on the *number of editorials* in which symbols appeared, not the frequency with which the symbol itself appeared.

Data from 19,553 editorials were used to trace changing foci of attention and attitude, as indexed by key symbols, for the sixty-year period. Among many findings were the following:

Symbols of representative government are used where the practice is under dispute, not where it is an accepted part of the traditions. (Pool, 1952b, p. 72)

Hostility to the outside world . . . seems to be very much of a function of insecurity. Those nations which have at any given moment dominated the world scene have generally said little that was adverse to “prestige papers” in the other power. The insecure or unsatisfied powers, on the other hand, have generally had editorials full of hostile judgments of foreign states. (Pool, 1951, p. 62)

Two main trends in the modern world are: (1) a shift in the center of attention, in which traditional liberalism is being replaced by proletarian doctrines, and (2) a growing threat of war and a corresponding increase of nationalism and militarism. (Pool, 1952 p. 84)

One important area of inquiry left relatively untouched by the RADIR studies is that of differences and trends in meaning, beyond approval or disapproval, of such symbols as "democracy" or "peace" (Stedman, 1951; Rainey, 1966). This is an area in which a variety of content-analysis methods might usefully be applied to supplement more traditional approaches to textual exegesis.

Other somewhat similar studies of political symbols have examined trends in Soviet May Day slogans (Yakobson and Lasswell, 1949), symbols used by the Communist International (Leites, 1949), and response of Communist propaganda to defeat in elections and strikes (Leites and Pool, 1949).

Various types of literature have been examined to reveal trends in religious values. Hamilton's (1942) examination of themes relating to faith, the nature of man, science, education, and social problems in published Protestant sermons demonstrated a significant change from social optimism to social pessimism during the decade prior to World War II. Forty-six inspirational best sellers published between 1875 and 1955 were analyzed in order to identify themes and trends in American popular religion (Schneider and Dornbusch, 1958). Constant themes during the period included the views that religion promotes success; that the individual can make changes beneficial to himself by religious means; and that man is inherently good. A number of other themes increased substantially in the more recent literature: religion is linked to national political aspirations, and religion promotes mental and physical health. On the assumption that individuals are motivated to purchase books which coincide with their own values (a premise well supported by independent evidence), and in view of the best seller status of this literature, inferences regarding trends in popular religious values can also be drawn from these findings.

These studies illustrate trend analyses in which a sample of relevant material is compared continuously over a period of time. In other cases the investigator may simply base his comparisons on two points in time. The validity of assertions identifying anti-intellectualism as an important American characteristic was tested by Hage (1959). Media analyzed were newspapers published during the 1828 Jackson-Adams and 1952 Eisenhower-Stevenson campaigns, in each of which an "intellectual" ran unsuccessfully against a military hero. Issues, candidates' traits, and their positions on the issues were recorded for two newspapers during each campaign. In 1952, the press was more prone to discuss the campaign in terms of issues, rather than personalities. Comments about Stevenson's intellectualism were predominantly favorable, although significant differences appeared between the "quality press" (favorable) and the "mass press" (unfavorable). The press of 1828 dealt with Adams' intellectualism and Jackson's unreflective traits strictly along partisan lines. In 1952, on the other hand, all segments of the press—mass and quality, pro-Stevenson and pro-Eisenhower—commented favorably on the latter's nonintellectual attributes. The results suggested considerable anti-intellectualism in the press during both campaigns. Whether these findings depict two isolated instances of anti-intellectualism or a persisting aspect of American culture cannot, of course, be determined solely from a study of this type.

Other trend analyses have examined various aspects of newspaper development, with heavy emphasis on American newspapers. Trends in general dailies, country weeklies, comics, and Presidential news have been described by Mot (1942), Tauber (1932), Barcus (1961), and Cornwell (1959), respectively. Considerable research has also been devoted to surveys of other sources, including movies (Dale, 1935), radio (Albig, 1956), television (Smythe, 1953), propaganda (Kris and Leites, 1953), popular heroes in magazines (Lowenthal, 1944; Winick 1963), and child-development materials (Ojemann *et al.*, 1948).

Trend inventories have varied widely in purpose and quality. Such studies can be useful for identifying major changes across long periods of time, and are relatively easy to undertake; on the other hand, surveys depending on gross categories often conceal more information of interest than they reveal. According to one source, "The most valuable use of studies of content . . . is in noting trends and changes in content. Systems of classification may be inadequate and unstandardized; nevertheless, if a system is used consistently over a time period valuable facts may appear" (quoted in Berelson, 1952, p. 29). This seems a dubious premise upon which to stake very much research effort. Certainly, little is to be gained by precise measurement if the system of categories employed is inadequate.

A special application of the trend study is the analysis of professional publications to index changing foci of scholarly interest. Articles in the *American Journal of Sociology* were classified by central topic to describe trends in the discipline (Becker, 1930, 1932; Shanas, 1945). As Berelson has pointed out, studies based solely on a few professional journals may distort the actual focus of research interest (1952, p. 34). Changes in classification of articles on the borderline of two disciplines might affect the results of studies covering a long time span. An additional problem is that of editorial policy. For example, articles on social reform accounted for 13 percent of all space in the *American Journal of Sociology* at the turn of the century, but no such article was published after 1935 (Shanas, 1945). Does this finding indicate changing professional interests, a shift in editorial taste, or perhaps establishment of a separate new journal for the specific purpose of presenting policy-oriented research?

Owing to a broader research design, a study of changing interests in American psychology over the half-century prior to 1938 is less vulnerable to this criticism. Articles appearing every tenth year, in 14 journals rated most "significant" by members of the profession, were coded for 32 categories reflecting the type of subjects used, topics, techniques of investigation, and modes of conceptualization. Each article was thus listed under a number of categories, usually ranging from five to 15 (Allport and Bruner, 1940). Use of more than a single journal and coding into multiple categories minimized potential distortion.

Other studies have examined trends in fields as general as physics (Rainoff, 1929), journalism (Tannenbaum and Greenberg, 1961b; Webb and Salancik, 1965), and botany (Stevens, 1932), and as specific as content analysis (Barcus, 1959) and Freud's dream theory (Herma, Kris, and Shor, 1943).

To relate known characteristics of sources to the messages they produce

An important aspect of social research is the testing of hypotheses of the form, "sources with characteristic *A* are likely to produce messages with attributes *w* and *x*, whereas those with characteristics *B* are likely to produce messages of types *y* and *z*." This formulation covers a broad spectrum of studies. Sources may be two individual authors or newspapers; or different media, such as radio and magazines; or communication sources in two or more different countries.

The relationship between ideological orientation of media and the nature of their news reporting has been an area of considerable research interest in Europe and the United States. As early as 1910 Max Weber urged (quoted in Krippendorff, 1966, p. 5):

... we will have to start measuring, plainly speaking, in a pedestrian way, with the scissors and the compass, how the contents of the newspapers has quantitatively shifted in the course of the last generation . . . between feuilleton and editorial, between editorial and news, between what is presented as news and what is no longer offered . . . and from the quantitative results we have to move toward qualitative ones. We have to pursue the style of presentation of the paper, the way in which similar problems are treated inside and outside the papers, the apparent expression of emotions in the papers.

More recently Gerbner has been among the most articulate proponents of the position that all analysis of the mass media must proceed with a sensitivity to the ideological framework of the media: "all editorial choice patterns in what and what not to make public (and in what proportion, with what emphasis, etc.) have an ideological basis and a political dimension rooted in the structural characteristics of the medium" (1964, p. 495). The general proposition that "all news are views" was tested by an examination of nine French newspapers of the political left, the right, and the commercial press. Reporting of a nonpolitical incident, the shooting of a schoolboy by a teacher, was subjected to a propositional analysis. On the basis of significant differences among newspaper types, the author concluded: "There is no fundamentally non-ideological, apolitical, non-partisan news gathering and reporting system" (Gerbner, 1964, p. 508).

Interest in the role of the press in electoral campaigns has been strong, especially since the Presidential elections of 1936 and 1940, in which the vast majority of the press supported Landon and Willkie in their campaigns against Roosevelt. Several studies have sought to determine whether editorial support is systematically related to other aspects of campaign coverage, such as the amount of space devoted to stories of each candidate. Evidence from such investigations is mixed, depending largely on the selection of newspapers and elections. Blumberg (1954) examined 35 American dailies during the 1952 campaign and found little evidence of bias in news coverage. This finding was supported by Markham and Stempel (1957). Studies of 15 "prestige" dailies during the 1960 and 1964 campaigns indicated that "as a group they gave the Democratic and Republican campaigns virtually equal amount of space in their news columns" (Stempel, 1961, p. 157; Stempel, 1965). But analyses of dailies in Florida and California revealed that endorsed political candidates received better coverage (Kobre, 1953; Batlin, 1954).

A serious limitation of investigations based on space measures is that they tap only a single dimension of bias. Some rough equality of space allocation may be a necessary condition of unbiased coverage, but probably it is not sufficient. During the 1940 campaign the press and radio focused on Roosevelt by a margin of 3-2, but favored Willkie by better than 2-1 (Lazarsfeld, Berelson, and Gaudet 1944, p. 117). The more subtle, and probably more important, methods of slanting news have received less attention than measures of space. One exception is an analysis of eight major daily newspapers during the 1952 campaign. Newspapers were rated on eighteen indices, including size and tone of headline, placement of stories, number of biased remarks, number of pictures, and total column inches of stories on various pages. These measurements gave strong indications of systematic bias in favor of the endorsed candidate (Klein and Maccoby, 1954).

A similar finding emerged from a study of the British press, which concluded that, with the exception of *The Times*, "There could be no doubt in the reader's mind as to which side the different newspapers supported. News and comment were inextricably mixed in the 'news' reports and special articles" (Royal Commission on the Press, 1949, p. 359).

Differential coverage of "civil rights" stories has also been related to various characteristics of newspapers, including geographical location, ownership (Negro, white), and political orientation. In general, the findings have supported hypotheses of systematic quantitative and qualitative differences in news coverage (Broom and Reece, 1955; Carter, 1957; Breed, 1958).

The hypothesis that local newspaper competition is related to more adequate news coverage has received some attention. A comparison of 97 newspapers revealed no significant difference in allocation of nonadvertising space related to the presence or absence of competition. A second test of 260 newspapers matched for circulation and other characteristics replicated the analysis, and the data again rejected the hypotheses that competitive status has any significant bearing on allocation of news space (Nixon and Jones, 1956). Similar conclusions were reached by Willoughby (1955, p. 204). Adequacy of news coverage during the 1960 Presidential campaign was related to various characteristics of newspapers in a survey of 90 dailies (Danielson and Adams, 1961). From this research a "news potential index" was developed which, on the basis of five newspaper attributes, predicts the adequacy of news coverage.

Studies comparing the content of two or more mass media have been relatively infrequent. In this respect development of new modes of communication has had less impact on content-analysis research than might be expected. This can be attributed in part to the problem of devising coding units, other than simple time and space measures, which can be compared. A strong bias toward research on the more easily accessible printed materials has also restricted cross-media comparisons and limited the kinds of statements that can be made about the media.

The question of what happens to a book when it is adapted into a movie was examined by Asheim (1950). An "index of deviation," based on the proportion of space devoted to main story lines and subplots in 24 books and movies, revealed

that fidelity in adaptation ranged from 89 percent (*Pride and Prejudice*) to 38 percent (*The House of Seven Gables*). Book and film endings were classified as "happy" or "unhappy"; in 21 cases the type of ending was accurately portrayed in the film, although details of the action were altered in three of them. The study also revealed that chronological sequence was followed more closely in the film than in the book, and that the incidence of violence, brutality, and sadism was reduced in the film.

Mental health content of the mass media has been the subject of a number of cross-media studies (Taylor, 1957; Nunnally, 1957; Gerbner and Tannenbaum, 1960). Space allocations differed considerably, broadcasting media (radio and television) carrying more relevant material than printed media (magazines and newspapers). However, the content of assertions about mental health was almost identical across all media. The three most frequently appearing themes were: the mentally ill look and act differently; mental health problems originate in environmental stresses; and the problem is a serious one (Taylor, 1957).

Content analysis has frequently been used on documents produced by political action groups. Comparative analysis of Communist publications was used to develop a model against which perceptions and experiences of former party members (determined by interviews) could be compared (see Table 2). Categories incorporating qualities of the "ideal communist" were used to code a Communist classic, publications of a party in power (Soviet Union), and of a weak party (United States) (Almond, 1954, p. 77).

TABLE 2

QUALITIES ATTRIBUTED TO THE SELF IN STALIN'S "HISTORY," THE COMINFORM PERIODICAL, AND THE "DAILY WORKER"

Qualities	History	Cominform	Daily Worker
Goal Qualities	6 %	8 %	26 %
Esoteric	4	0.5	
Exoteric	2	7.5	26
Tactical Qualities	94	92	74
Militance	28	25	35
Rationality	23	10	2
Organization	16	9	3
Leadership	13	19	11
Activism	3	7	6
Uniqueness	3	6	1
Dedication	3	7	13
Confidence	5	9	3
Total percent	100	100	100

The American business community (Lane, 1951; Bernstein, 1953), right-wing organizations (Wilcox, 1962; Abcarian and Stanage, 1965), and various groups of lobbyists (McPherson, 1964) are among other groups whose publications have been analyzed.

Content analysis has also been used on a variety of materials to discover international differences in the content of communication. A comparative study of newspapers in 17 countries during a seven-day period in March 1951 was conducted under auspices of UNESCO (Kayser, 1953). The morning newspaper with the largest circulation in each country was analyzed for a period of time selected in advance of actual publication. Comparisons were made of both format and content: the front page, space allocation, origin of news, and coverage of specific events during the week. A supplementary analysis compared four "prestige" papers published in Moscow, Paris, London, and New York. Other cross-national studies have compared aspects of the press in the Middle East (Abu-Lughod 1962), Australia-New Zealand (Budd, 1964), and the United States-Great Britain (Hart, 1965, 1966).

An examination of fifteen "prestige" newspapers published on November 2, 1956, during the dual crises in Hungary and Suez, revealed differences in coverage of these events which were ascribed to the degree of involvement in one event or the other, or to instrumental handling of the news. News of the Suez invasion crowded Hungary out of these newspapers; even in the West, "attention was so overwhelmingly on Suez that the full significance of what was happening in Hungary was never made clear" (Schramm, 1959, p. 138). A less comprehensive study compared Indian and German coverage of the concurrent crises on the Sino-Indian border and in Cuba during October 1962 (Roat, 1963).

Cross-national content analysis has not been confined to printed media. Wolfenstein and Leites (1950), in a study of plot configurations of British, French, and American movies, found that British movies emphasized that danger lies in the nature of man himself, especially in his impulses of destructiveness; many of the themes of Shakespearean drama were preserved, although cloaked in a modern idiom. French movie plots revealed human wishes opposed by the nature of life itself which, in the end, defeats all, including the virtuous, and both old and young suffered inevitable disappointments in love. In American movies, hazards of life were found in the situation, rather than in the nature of the individual or life itself. Winning was stressed, although not always achieved easily, and disappointment in love was denied.

Two studies compared the literature and songs of youth groups in Nazi Germany and the United States. Youth literature in the two countries yielded nearly 1000 expressions of various organizational goals (Lewin, 1947). By frequency count, the German literature placed significantly greater stress on national loyalty, national identification, and determination, whereas American Boy Scout materials emphasized altruism, religion, and creativity. A parallel study of children's songbooks yielded similar results (Sebald, 1962). German sources stressed national loyalty, obedience, and heroic death, and paid less attention to the beauty of nature, play, and Christianity. Content analysis of such materials can reveal

important international differences at a specified point in time, but further inferences, unsupported by independent data, are often open to question. A case in point is Sebald's generalization, based solely on Nazi songbooks issued in 1940, that the modal character of Germans is basically authoritarian.

Content analysis has also been used to examine cross-national or cross-cultural differences in the content of television programming (Gardiner, 1962), magazines (Ginglinger, 1955), textbooks (Walworth, 1938), folktales (Colby, Collier, and Postal, 1963), sermons (Parker, Barry, and Smythe, 1955), and magazine photographs (Wayne, 1956).

To audit communication content against standards

Content data have often been compared to *a priori* standards. Early content-analysis research in the United States was stimulated by concern over the spread of yellow journalism; studies of New York newspapers by Speed (1893) and Matthews (1910) are examples. This research was almost wholly devoted to measuring space allocation for various subject-matter categories, and much of it has been justly criticized for subjective and arbitrary procedures; Matthews, for instance, aggregated his categories into four major classes: "demoralizing," "unwholesome," "trivial," and "worthwhile." Interest in standards of the mass media has been sustained, however, and many serious technical problems of the early studies have been resolved.

The standard of "social norms" was used in a study of goals and goal achievement as portrayed in television programs for children, adult, and mixed children-adult audiences (Larson, Gray, and Fortis, 1963). Each act by a character which was identifiably connected with a specific goal was coded. Goals were classified into categories such as power and prestige, property, self-preservation, affection, sentiments, and psychological goals. Means to achieve the goals were coded as socially approved (legal) or socially disapproved (nonlegal, violent, or escape). Results revealed that means with the *least* likelihood of success were those classified as socially approved. There was no essential difference in this respect between those programs which were viewed primarily by adults and those designed for younger audiences.

A study of three major American news magazines—*Time*, *Newsweek*, and *U.S. News & World Report*—against the standard of "responsibility in mass communication" indicated that the "neatly reconstructed picture of the world" they present is often biased, distorted, or factually false (Bagdikian, 1959). While apparently strong evidence of systematic bias was presented, the absence of explicit coding categories or sampling methods raises some questions about the findings. A rigorous study of the same news magazines during the period of the 1960 conventions revealed that Republican candidates as a group received only slightly more favorable treatment than their Democratic counterparts (Westley *et al.*, 1963). But when the data were reexamined according to individual candidates, each magazine was found to have treated the conservative candidates (Nixon and Johnson) more favorably than the liberal ones (Rockefeller and Kennedy).

An analysis of six types of bias (attribution bias, adjective bias, adverbial bias, contextual bias, outright opinion, and photographic bias) revealed that *Time* used each technique extensively to describe recent American Presidents. *Time's* presentation of Truman was totally negative, Eisenhower was depicted in an unambiguously favorable light, and only Kennedy was described in somewhat balanced terms (Merrill, 1965). "Repeated distortion and misinformation" were also discovered in public affairs articles in *Reader's Digest* (Christenson, 1964).

Ash (1948) examined 50 periodicals over a period of 10 months to determine whether the public was given a fair opportunity to learn both sides of the controversial Taft-Hartley Labor act. The analysis uncovered almost unanimous support for the act; moreover, those few anti-Taft-Hartley items which did appear were largely confined to low-circulation periodicals. High-circulation news magazines failed to print any items opposing the act, and the general magazines were similarly one-sided. When item counts were adjusted for circulation, 55.7 percent of the material favored the act and 1.8 percent opposed it, the remainder being neutral. However, no attempt was made to test the relevance or justice of the arguments, pro and con.

Another aspect of distortion was studied by Cony, who tested the hypothesis that "Newspapers emphasize conflict to the extent that reality is twisted out of shape and a false picture of society as a jungle is presented to the reader" (1953 p. 15). The data revealed conspicuous differences among a sample of five newspapers; but, although 1952 was a time of considerable conflict, a substantial amount of space (31 to 46 percent) was devoted to reporting cooperative behavior.

The most common weakness of studies using an *a priori* standard stems from the absence of a clearly defined basis for judgment, rather than from technical problems of applying content analysis to the data. What constitutes "adequacy" in the coverage of local news? Or, how close to "equal time/space" must the media come to be considered "fair"? It is doubtful whether unbiased reporting of every controversial situation calls for equal presentation or evaluation, as is sometimes assumed by indices of bias. Often the investigator's own values serve as the standard of comparison, but failure to define such terms explicitly make it difficult to interpret findings. In his study of country weeklies, for example Willey concluded that "The Connecticut weekly newspapers are deficient in the amount of local news material they print" (1926, p. 111). Yet his data revealed that space allocation to local news ranged from four percent to 92 percent, and more than half the newspapers devoted over 50 percent to local news.

One answer to the difficult problem of defining standards against which to audit sources is to make comparisons against other sources; that is, general norms for classes of communicators are developed inductively. The investigator may then rank sources along one or more dimensions. This technique has been used to rate radio commentators (Budlong, 1952) and reporting of specific events by magazines, newspapers, radio, television, and wire services (Rosi, 1964; Klappe and Glock, 1949; Sussman, 1945; Lang and Lang, 1955; Rucker, 1960). A more precise method, the technique of "successive approximations," was used to construct a "socialization-sensationalism" index. Six rough indicators were first

CHARACTERISTICS OF CONTENT: HOW

To analyze techniques of persuasion

Content analysis usually focuses on the substance (the "what" question) of messages. It has also been used to analyze form or style (the "how" question). For the past three decades, and particularly during World War II, considerable research has focused on propaganda, "the manipulation of symbols as a means of influencing attitudes on controversial matters" (Lasswell, 1942, p. 106). Often the purpose has been to infer intentions of communicators from propaganda content, a type of analysis to be discussed later. The remaining research has aimed at developing a theory of form, style, and structure of persuasive communication.

A pioneering study in this area was Lasswell's analysis of propaganda techniques during World War I, in which four major objectives of propaganda and appropriate techniques of appeal for each goal were identified: (1) to mobilize hatred against the enemy; (2) to preserve the friendship of allies; (3) to preserve the friendship and, if possible, to procure the cooperation of neutrals; (4) to demoralize the enemy (1927, p. 195). Lasswell concluded that, although all four themes were present in the propaganda of every nation, they were applied with varying degrees of success; much German propaganda turned out to "boomerang," partly because of the lasting impression that Germany was the aggressor, partly because of the ineptness of the appeals. On the other hand, British propagandists were successful in picturing humanitarian war aims, and the French were able to portray the Germans in satanic terms, such as "Hun," and "Boche" (1927, pp. 196-199). More quantitative methods were used to examine the organization, media, techniques, and symbols of Communist propaganda in Chicago during the depression of the 1930's (Lasswell and Blumenstock, 1939). The Lasswellian influence in propaganda studies continues to the present, as is evident in two recent books (Dovring, 1959; Barghoorn, 1964).

A different conception of propaganda, stressing omissions and selectivity in presentation of factual materials rather than use of demagogical tricks or misuse of logic, was tested by examining arguments relating to the acquisition of nuclear weapons in Swedish newspaper editorials (Ohlström, 1966).

Several sets of categories for describing and analyzing various aspects of propaganda have been proposed. One scheme, developed at the Institute of Propaganda Analysis, enumerates content (name calling, testimonial, bandwagon, etc.) and strategic (stalling, scapegoating, etc.) techniques which have been identified in propaganda (Lee, 1952, pp. 42-79, 210-234).

A somewhat different set of categories emerged from a comparative study of British and German radio broadcasts to the United States during 1940 (Bruner, 1941). A tentative list identified nine dimensions for describing propaganda: dissolvent-unifying, negative-positive, temporal, personal-impersonal, stratified-homogeneous, authoritative-casual, colloquiality, immediate-remote, and repetitiousness.

"Value analysis," a set of categories for studying personality from written materials, was used to examine the propaganda style of Hitler and Roosevelt (White, 1949; for a detailed explanation of value analysis, see White, 1951). A number of similarities were found: both Hitler and Roosevelt stressed traditional grandeur, and both often used black-white dichotomies. Hitler also appears to have used an indirect approach to the preparation of the German people for war by emphasizing the theme of persecution by outsiders. The same method and categories were found useful for distinguishing writings by political figures of various ideological persuasions: Khrushchev, Stalin, Hitler, Mussolini, Goldwater, Hoover, Churchill, Kennedy, and Franklin D. Roosevelt (Eckhardt, 1965). Other propagandists whose techniques have been examined by means of content analysis include Father Coughlin (Lee and Lee, 1939) and Gerald L. K. Smith (Janowitz, 1944).

A study of the picture magazines *USSR* and *American Illustrated*, which are produced by the Soviet Union and the United States for readers of the other country, revealed how persuasive literature is often framed in the value context of the audience. Both the American and Soviet magazines put greater emphasis on values often attributed to the other nation; that is, *USSR* emphasized such aspects of Soviet life as industrial growth and a high standard of living, whereas *American Illustrated* stressed the cultural and esthetic interests of its citizen (Garver, 1961).

Although the most pervasive form of persuasive communication—advertising—has received comparatively little attention from content analysts, research has not been limited to official governmental propaganda. For example, campaign biographies, a form of persuasive literature which appears on the American political scene quadrennially, have been analyzed. The basic theme, to fashion an image of the ideal citizen of the Republic, has remained the same since 1824. Moreover, "rival candidates appear in campaign biographies to be as alike as Tweedledum and Tweedledee" (Brown, 1960). Considerable content-analysis research has also focused on public letters, such as those written to congressmen (Wyant and Herzog, 1941), to various newspapers in the United States (Foster and Friedrich, 1937; Toch, Deutsch, and Wilkins, 1960), the Soviet Union (Inkeles and Geiger, 1952, 1953), and Communist China (Wang, 1955), and to a magazine by radical right-wingers (McEvoy, 1966).

The most evident weakness of propaganda analysis has been the absence of systematic research to relate categories of appeal, techniques, and dimensions (and combinations of these) to effects. What types of appeals are most effective? Under what circumstances? For which subject matter categories? One exception is a detailed investigation of Kate Smith's war bond drive. Content analysis was used to identify characteristics of her appeals which might be expected to elicit particular responses from the audience. The validity of inferences based on the content data was then checked by interviewing the audience (Merton, 1946). But in the main, questions about how technique and content of appeals are related to effects have remained unanswered.

To analyze style

Studies of style have differed widely in method and have ranged from investigations of a single author to analyses of an entire language. Word counts have yielded concordances of individual authors, including Yeats (Parrish and Painter, 1963) and Matthew Arnold (Parrish, 1959), and have been used to describe continuities and discontinuities in the Russian language (Josselson, 1953). Characteristics of style which distinguish good paintings from mediocre ones and best-selling novels from those less successful financially have been investigated (Gordon, 1952; Harvey, 1953). And relative frequencies of various parts of speech—for example, verb-noun-adjective ratios—were used to describe constant and varying characteristics of poetic style across five centuries (Miles, 1951).

Generalizations about qualitative features of literature have been tested by content analysis. The metaphorical qualities of Jane Austen's *Persuasion*, Emily Brontë's *Wuthering Heights*, and George Eliot's *Middlemarch* were examined to illustrate the thesis that "metaphorical language reveals to us the character of any imaginative work . . . more tellingly perhaps than any other element" (Schorer, 1949, p. 560). Skinner analyzed the pattern of alliteration—the appearance of two or more syllables beginning with the same consonant near each other—in Shakespeare's sonnets. After adjustment for the number of occurrences deriving from repetition of the same words, the frequency of alliteration did not differ significantly from chance; "so far as this aspect of poetry is concerned, Shakespeare might as well have drawn his words out of a hat" (Skinner, 1939, p. 191). A related aspect of style, in which the character rather than the word or phrase serves as the recording unit, is letter redundancy. Paisley (1966) content-analyzed 39 samples from English translations of Greek texts covering three time periods, 18 authors, and nine topics. Letter redundancy was shown to vary systematically with authorship, topic, structure, and time of composition.

Political rhetoric, especially that of American Presidents or Presidential candidates, has been a favorite subject for study. One approach is illustrated by a study of Woodrow Wilson's speeches (Runion, 1936). Categories were developed around grammatical aspects of discourse—sentence length, sentence structures, and figures of speech. A more revealing examination of political rhetoric, a study of broadcast addresses by Eisenhower and Stevenson during the 1956 campaign, related substantive and stylistic categories (Knepprath, 1962). Styles of the two candidates were compared for subject matter, form of reasoned discourse, use of "loaded terms," and types of motive appeals.

CHARACTERISTICS OF CONTENT: TO WHOM

To relate known characteristics of the audience to messages produced for them

The proposition that communicators tend to cast their messages in the idiom of the intended audience has been tested occasionally. Such studies have tended to center on messages produced to change attitudes. An examination of two mass circulation magazines, *Life* and *Ebony*, supported the hypothesis that differences

in advertising are based on the socioeconomic levels of the two audiences (Ber man, 1963). When the offered product or service was predicated on the Negro present lower status, the advertising was in a form appropriate to that status. On the other hand, when the advertised items reflected the status to which the Negro was thought to aspire, there was little essential difference in the two magazines except for the substitution of Negro models in *Ebony*, two-thirds of the advertisers selling to middle-class markets used identical ads in the two magazines.

Albrecht (1956) analyzed short stories in large-circulation magazines with lower (*True Story*, *True Confessions*), middle (*American*, *Saturday Evening Post*) and upper (*Atlantic*, *New Yorker*) level readers, recording the distribution and evaluation of 10 "family values." Stories produced for each level strongly supported American family norms as a group, the least deviation being found in magazines for middle level readers. The degree of approval of some specific values, however, differed significantly from level to level.

Such data may lend themselves to three different interpretations: that authors *write differently* for dissimilar audiences, that the literature *reflects* basic value differences of the audiences, or that such materials *shape* the values and predispositions of the audience. These are not mutually exclusive aspects of the communication process, but rather, represent three types of inferences which have been drawn from content data. The first interpretation presents the fewest problems of validity and can be made directly from attributes of the content. The second explanation is a variation of the theme, often debated with respect to the mass media, that producers of communication are only giving audiences "what they want." This may or may not be true, and generally can be confirmed only by means other than content analysis. The third inference is the most tenuous unless supported by data other than the description of the content.

To describe patterns of communication

How are patterns of communication affected by situational or systemic changes? Analysis of documents written by leaders of the Dual Alliance and Triple Entente nations during the summer of 1914 indicated that as war approached, there was a significant increase in messages exchanged within alliances, with a concomitant decrease in intercoalition communication (Holsti, 1965b). Messages produced in 17 "Inter-Nation Simulations" revealed that after all members of an alliance had obtained nuclear weapons, the modal pattern of communication changed from a "wheel" configuration to an "all channels" pattern (Brody, 1963).

THE ANTECEDENTS OF CONTENT: WHY

Many definitions of content analysis explicitly exclude its use for purposes of inferring the antecedents or causes of content. Nevertheless, such inferences have often been drawn from content data. One major goal of propaganda analysis, for example, has been to make inferences about values, intentions, and strategy of communicators. And, as indicated earlier, research on the pragmatic dimension of communication has become a major aspect of content-analysis studies. The

problem, then, is no longer whether content analysts *should* make inferences concerning the cause of communication, but rather, given the trend of research, what steps the investigator can take to enhance confidence in the validity of his inferences. Studies selected for review in this section were chosen in part to illustrate a variety of research applications and in part to examine alternative approaches to the problem of validation.

To secure political and military intelligence

An important impetus to the development of content analysis was the large-scale propaganda research during World War II. Social scientists, many of whom later made significant theoretical and methodological contributions to content analysis, were engaged by the Federal Communications Commission, the Library of Congress, and the Justice Department to study these materials.

The most difficult problem, because of the constraints within which the propaganda analyst operates, is that of establishing criteria for inference. The FCC used both *direct* and *indirect* techniques. The first method operates from a "representational" model of communication; that is, the investigator assumes that words in the message are valid indicators, irrespective of circumstance. Inferences regarding intentions, expectations, and situational factors are drawn *directly* from attributes of propaganda, based on a past correlation of conditions or events and content characteristics. The direct method is illustrated by a study to determine the degree of collaboration between German and Italian propaganda agencies. From consistent differences in broadcasts originating in Rome and Berlin, analysts concluded that there was no collaboration. Evidence gathered after the war validated the inference (Berelson and De Grazia, 1947).

The single-step approach has been criticized for two deficiencies: past regularities are often based on a very few cases, and the method is insensitive to changes in propaganda strategy, which may render past correlations invalid.

An "instrumental" model of communication, in which it is assumed that the important aspect of the message consists in what it conveys, given context and circumstances, underlies the indirect method of inference (the representational and instrumental models are further discussed in Pool, 1959). The initial step in the indirect method is to establish the propaganda goal or strategy underlying the characteristics of content. A series of interconnected causal imputations is derived from this point (George, 1959a, p. 41):

Situational ← Elite ← Elite ← Elite intention ← Propaganda ← Content
actor ← estimate ← expectation ← or policy ← strategy

In the indirect method, then, the process of inference is broken up into a number of smaller steps. The relationship of the two methods of inference to one another is summarized by George (1959a, p. 43):

It is [recommended], first, that the direct approach be utilized in *hypothesis formation* or whatever it is worth (which depends upon the number and quality of the tentative, incompletely confirmed one-to-one correlations that are available) and that the indirect

method be utilized in the *assessment*, or testing, of the inferential hunches derived from the direct approach. Second, it is recommended that such transitions from the direct to the indirect approach be made quite deliberately, in full awareness that the indirect method requires logic-of-the-situation reasoning and the use of generalizations other than the one-to-one type of correlation between a content indicator and an aspect of the elite's political behavior or situational milieu.

Despite many difficulties facing analysts, documentary material on the Nazi conduct of the war indicated that FCC inferences were accurate in an impressive number of cases. For a two-month period (March–April 1943), 101 out of 111 inferences made by the German section were scored as correct. Of methodological interest is the finding that frequency and nonfrequency indicators were about equally successful (George, 1959a, pp. 264–266).

To analyze psychological traits of individuals

It is a widely held belief among social scientists that symbolic behavior of the individual can provide important psychological data about personality, values, intentions, and other characteristics of the communicator. Personal documents—defined as "any self-revealing record that intentionally or unintentionally yield information regarding the structure, dynamics, and functioning of the author's mental life"—may take many forms, ranging from a diary or intimate letters to autobiographies and speeches addressed to a wide audience (Allport, 1942, p. xii). The motives for producing personal documents may vary from psychotherapy to hopes of literary fame. Finally, investigators' purposes in analyzing personal documents have differed, despite the common goal of making inference about the communicator.

Written materials have been content-analyzed to infer personality traits of their authors. A study of Richard Wright's autobiography (White, 1947) was discussed earlier. The letters of an Irishwoman, Jenny Gove Masterson, were subjected to "personal structure analysis," a system developed to aid the clinician "as a supplement to his more or less brilliant insight, a technique offering evaluation and analysis which will have the virtue of objectivity and will also reveal aspects of the material that may have eluded his scrutiny" (Baldwin, 1942, p. 163). The procedure was grounded on the assumptions that frequency of an item in the case material is a measure of its importance in the personality. From a table of frequencies and correspondences, three well-integrated but independent ideational clusters were isolated. These revolved around Jenny's attitudes toward self and son, jobs, and death (the text of the letters can be found in Allport, 1946). Baldwin's study has been replicated using computer analysis (Paige, 1966).

Although critical of Baldwin's statistical analysis, Andrews and Muhlhan employed a modified version of personal structure analysis to study congruent idea patterns in the personal diary of a young girl (1943). As in Baldwin's study, independent checks were used to test the validity of content-analysis data, with satisfactory results. The method of "contingency analysis" employed in both of these studies has been further developed by Osgood (1959).

A number of categories have been used to analyze themes, plots, and characters of novels to infer psychological traits of their authors. As in other studies of this type, the analyst assumes that key psychological traits of the writer must manifest themselves in his writing. Novels by D. H. Lawrence (McCurdy, 1939), the Brontë sisters (McCurdy, 1947), Charles Kingsley (Deutsch, 1947), Dostoevsky (Kanzer, 1948), and Knut Hamsun (Lowenthal, 1949) have been analyzed, as have Shakespeare's plays (McCurdy, 1953). Studies of this nature have not been limited to fictional materials. The content of Justice Robert Jackson's legal opinions, for example, were analyzed and related to other aspects of his behavior in an effort to reconstruct continuities and changes in his legal philosophy (Schubert, 1965).

Unlike descriptive studies of literary style, such as those discussed earlier, problems arising in the analysis of literature for the purpose of inferring psychological traits are not confined to those of developing adequate analytic categories. A number of methods have been used to validate content data. Investigation of the author's biography may provide at least a partial check on the validity of inferences drawn from the data; that is, are events and experiences in the writer's life consistent with psychological inferences drawn from content data? This method was applied in the studies of the Brontës, Dostoevsky, Lawrence, Shakespeare, and Jackson.

An interesting variation of the validity problem, validation by prediction, is illustrated in Lowenthal's 1937 study of Knut Hamsun. Hamsun's novels revealed themes consistent with fascism: stress on race and natural community, pantheism, and reduction of women to reproductive functions. From these data Lowenthal concluded that Hamsun was intrinsically a fascist, a conclusion borne out a few years later by Hamsun's collaboration with the Nazis during World War II (1949).

A very different kind of personal document, political rhetoric, has been analyzed to infer personality traits of the speaker from logical and cognitive characteristics of his verbal production (Shneidman, 1963). The text is first coded into two category sets. *Idiosyncrasies of reasoning* include 32 categories consisting of idiosyncrasies of relevance, idiosyncrasies of meaning, arguments containing suppressed premises or conclusions, idiosyncrasies of logical structure, and idiosyncrasies of logical interrelations. *Cognitive maneuvers* consist of 65 styles of thought development; for example, to switch from normative to descriptive mode, or to render another's argument weaker or stronger by paraphrase.

To illustrate the method, Shneidman (1961, 1963) examined the logical styles of Kennedy and Nixon on their first two television debates, and that of Khrushchev in speeches delivered after the collapse of the Paris "Summit" Conference and at the United Nations. A partial comparison of "idiosyncrasies of reasoning" and "cognitive maneuvers" in the rhetoric of Kennedy, Nixon, and Khrushchev appears in Table 4.

The second step in the analysis is to construct, for each idiosyncrasy of reasoning, the logical conditions under which idiosyncrasy is controverted or canceled or, to use the author's term, *contralogic*. Inferences regarding psychological characteristics of the communicator are then drawn from the contralogic as illustrated in Table 5 (Shneidman, 1961, p. 22).

TABLE 4

IDIOSYNCRASIES OF REASONING AND COGNITIVE MANEUVERS IN THE RHETORIC OF JOHN F. KENNEDY, RICHARD M. NIXON, AND NIKITA KHRUSHCHEV

Stylistic features	Kennedy	Nixon	Khrushchev
<i>Idiosyncrasies of reasoning:</i>			
Irrelevant premise	8.7%	4.9%	2.2%
Argumentum ad populum	3.4	12.0	22.6
Complex question	0.0	1.9	21.5
Derogation	0.9	4.9	11.8
Stranded predicate	6.6	7.1	5.4
Truth-type confusion	2.2	6.4	0.0
<i>Cognitive maneuvers:</i>			
To enlarge or elaborate the proceeding	7.9%	6.0%	0.6%
To smuggle debatable point into alien context	5.8	8.4	0.0
To be irrelevant	7.3	9.6	16.8
To allege but not substantiate	4.4	6.3	4.6
To introduce new notion	0.0	0.0	2.3

Shneidman's inference regarding the personalities of Kennedy, Nixon, and Khrushchev appears to have considerable face validity. For example, Khrushchev is characterized in these terms (1961, pp. 61-62):

He feels that others are prone to misunderstand his position and yet he desires acceptance and will even sacrifice other needs or ends to achieve it. He is moody and needful of approval. But with his pessimism about resolving differences, he enjoys conflict and struggle, as much for its own sake as a means to an end. . . . He trusts his own instincts his "natural feel" for things. He is painstaking in certain areas, but in general is impatient and suspicious of detail or subtlety.

Pending considerable further research, the psychological correlates of logic styles are only working hypotheses. Nevertheless, as a method of studying style this technique represents a substantially more sophisticated approach than earlier attempts to analyze political discourse through content analysis (Hayworth, 1959; McDiarmid, 1937; Runion, 1936).

The utility of content analysis in psychiatry stems from the view that "psychiatric disorders, regardless of their etiology, are ultimately manifest as disorders of social communication" (Jaffe, 1966, p. 689). Sound recording of psychotherapeutic interviews has opened an entirely new field and a vast source of personal documents for analysis.

<p><i>Column I: Idiologic</i></p> <p>Samples of specific ways of thinking, in terms of idiosyncrasies of logic</p> <p><i>Example</i> (from Bleuler, von Domanus, and Ariety): "Switzerland loves freedom; I love freedom; therefore I am Switzerland." <i>Fallacy</i>: identification in terms of attributes of the predicate (or undistributed middle).</p> <p><i>Example</i> (from Binswanger's "The Case of Ellen West"): At 16 her motto is <i>Aut Caesar aut nihil</i> ("Either the greatest or nothing"). Before her death she says, "If I cannot remain young, beautiful, and thin, then rather—nothingness." <i>Fallacy</i>: dichotomization.</p> <p><i>Example</i> (from Wertheimer): Question to the young Gauss—"What is the sum of $1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10$?" <i>Response</i>: correct answer (55) in amazingly short time. <i>Fallacy</i>: none.</p>	<p><i>Column II: Contralogic</i></p> <p>Logical conditions under which the logical idiosyncrasy (column I) is con- verted or annulled</p> <p>If one supplies the implicit premise that Switzerland is the only member of the class of freedom lovers (and if I loved freedom), then it would follow, without logical error, that I would have to be Switzerland.</p> <p>If one supplies the implicit assumption that the aspects of the universe under discussion are organized in a binary manner, then it follows logically that the alternative to leading a perfect life is death.</p> <p>One can do the task correctly and quickly if one sees (makes the assumption) that (a) there are pairs of numbers ($1 + 10, 2 + 9, \text{etc.}$) each of which totals 11; (b) there are 5 such pairs; (c) 5 times 11 equals 55. The principle is represented by the formula $(n + 1) (n/2)$.</p>	<p><i>Column III: Psychologic</i></p> <p>Psychological state consistent with the logical condition (II) which would annul the idiosyncrasy (I)</p> <p>The reasoning reflects a psychological state in which the range of attention is constricted and narrowed (to one member of a class).</p> <p><i>Psychological symptoms</i>: intense concentration, oblivion to ordinary stimuli, hypesthesia, acute withdrawal, and, at its extreme, catatonia.</p> <p>The reasoning reflects a psychological state in which the basic identifications are confused, the basic attitudes toward self-control are ambivalent, and there is a general polarized approach to concepts and people, with unhappy vacillations between the extremes of the dichotomized polarities.</p> <p><i>Psychological symptoms</i>: obsessions, abulia, fixed ideas, anxiety, general impotence, rumination, suicide.</p> <p>The reasoning reflects a psychological state in which there is a high intelligence, active curiosity, open mind, freedom to explore intellectually, inquiring mind that is original but not destructive, freedom from fear of teacher (or authority generally? or father?) The psychological label for this state is "genius."</p>

Much research has been directed toward developing and validating measures for diagnosis and for evaluation of psychotherapy. A partial listing of such indices follows.

The *type-token ratio* (TTR) measures variability in the communicator's working vocabulary. A score is based on the number of different words found in samples of standard length: 100, 200, 500, or 1000 words (Johnson, 1944). The hypothesis that speech variability increases with successful therapy has generally been supported. A test of 12 schizophrenics and 12 college freshmen revealed that the mean TTR for schizophrenic patients was significantly lower; also, the stability of their TTR scores in successive samples was significantly lower than that of freshmen (Mann, 1944; Fairbanks, 1944). In addition, the language of the patients was characterized by a more negative tone, preoccupation with the past and more frequent self-references. Roshal (1953) tested the hypothesis that language variability, as measured by TTR scores for both 100 and 200-word samples, is related to adjustment. Difference in TTR's between the first and final interview were significantly higher for the group in which therapy was judged to be "more successful using both measures." Further clinical applications of the TTR are described in Gottschalk (1961).

The *adjective-verb* ratio scores the number of adjectives per 100 verbs (Bode, 1940). This measure has been found to differentiate the language structure of "normal" subjects from that of schizophrenics. Normal subjects have more adjectives per verb, as well as more adjectives per noun (Mann, 1944).

The *discomfort-relief quotient* (DRQ), derived from learning theory, measures the amount of drive borne by the client (Dollard and Mowrer, 1947). The DRQ is computed by dividing the number of discomfort (drive) words by the total number of discomfort and comfort (relief) words. Tests with the DRQ have produced rather mixed results. In some cases it has been significantly correlated with measures such as palmar sweating, but in others no relationship with success in therapy was noted.

The *positive-negative-ambivalent quotient* (PNAvQ) resembles the DRQ except that, unlike the DRQ, only the patient's self-evaluations are scored (Rain, 1948). DRQ and PNAvQ scores were computed for 17 interviews, with strong evidence that the two scores are highly correlated (Kauffman and Raimy, 1949). However, this finding was not confirmed by a second study.

The hypothesis that there is a positive correlation between the extent to which an individual expresses acceptance of and respect for himself and the degree to which he expresses these feelings toward others has been investigated (although not with the PNAvQ). Although the samples studied were small (10 subjects), in both cases the data supported the hypothesis (Sheerer, 1949; Stock, 1949).

A measure of *defensiveness* was developed on the premise that defensiveness follows perceptions of threat, an experience perceived to be inconsistent with value or concept of self or the environment (Haigh, 1949, p. 181). A defensive reaction is one in which perceptions are distorted to reduce awareness of perceived incongruence, and is inferred from such client responses as denials, rationalizations, and projections.

Scales for thematic analysis of *hostility* and *anxiety* found in verbal samples have been developed. One is a weighted scale for measuring three types of hostility classified on the basis of direction—outward, ambivalent, and inward (Gottschalk, Gleser, and Springer, 1963). A second scale is based on a system of weighted scoring for six categories of anxiety—death, mutilation, separation, guilt, shame, and nonspecific anxiety (Gleser, Gottschalk, and Springer, 1961).

Lasswell's general-purpose system classifies responses according to communicator (self, interviewer, other), direction (pro or anti), and object of statement (self, interviewer, other) (Lasswell, 1938).

A measure of *speech disturbance* has been found to correlate significantly with anxiety and conflict in therapeutic interviews. The ratio is calculated by dividing the number of speech disturbances by the number of words in the sample (Mahl, 1959).

On the theory that "motivation and psychological conflict play significant roles in shaping and distorting language responses," a general approach to the study of normal and pathological speech by word associations has been developed (Laffal, 1965, p. x). The subject's entire vocabulary (rather than only certain types of words, as in the DQR, PNAVQ, and adjective-verb ratio), is coded into 114 categories for which extensive definitions and scoring rules are presented.

Content-analysis systems have also been developed for systematic examination of Rorschach (Elizur, 1949; Lindner, 1950) and Thematic Apperception Tests (Shneidman, 1951; Hafner and Kaplan, 1960). Research relating other language categories to psychotherapy is reviewed in Mowrer (1953), Snyder (1953), Auld and Murray (1955), and Mardsen (1965).

Content-analysis research in psychotherapy has yielded somewhat mixed results. In some cases, contradictory findings have been obtained using the same measure with different samples of subjects. In others, the small samples of subjects have tended to raise questions about the generality of the findings. The primary difficulty has often rested not with the analytical scheme for scoring responses, but with the absence of clearly defined criteria of "success" (Auld and Murray, 1955, p. 389). This is a restatement of the recurring problem in using content data to make inferences about communicators and audiences: content data are not self-validating, and clear measures of the dependent variables that they are intended to index are often lacking. Nevertheless, the psycholinguistic approach to therapeutic materials appears to offer a fertile area for research. Rarely has the richness of communication been so dramatically illustrated as in an extensive linguistic and paralinguistic content analysis limited to the first five minutes of a psychiatric interview (Pittenger, Hockett, and Danehy, 1960). Some of the possibilities in this area are also suggested in an extensive, multimeasure psycholinguistic analysis of two interviews. Minute-by-minute readings on heart rate and skin temperature were related to a number of verbal measures, including the TTR, Speech Disturbance Ratio, and rating scales for anxiety and hostility (Gottschalk, 1961).

Use of language measures for purposes other than psychotherapy can be illustrated by a study comparing genuine suicide notes with ordinary letters and simulated suicide notes (Osgood and Walker, 1959). Among many content-

analysis methods used were the TTR, DRQ, and adjective-adverb/noun-verb measures, all of which differentiated real suicide notes from the letters. Only the latter measure, however, was successful in distinguishing real from simulated notes. Content analyses of real and simulated suicide notes are also reported in Gottschalk and Gleser (1960) and Stone and Hunt (1963).

Several techniques for scoring verbal interaction among group members have been developed. Interaction Process Analysis uses 12 content categories for classifying responses in six problem areas: communication, evaluation, control, decision, tension reduction, and reintegration (Bales, 1950). Sign Process Analysis is another technique for recording systematically what is said in the course of group interaction. The scoring scheme abstracts and records the distribution of positive, negative, and neutral assertions about objects, which are grouped into subsets according to locus (internal or external), sex, and status (Mills, 1964, p. 103). The Leary content-analysis approach (1957, p. 65) identifies eight interpersonal processes: managerial-autocratic, competitive-narcissistic, aggressive-sadistic, rebellious-distrustful, self-effacing-masochistic, docile-dependent, cooperative-overconventional, and responsible-hypernormal. Although developed for the diagnosis of personality, the Leary System has been adapted for use in simulation studies (Brody, 1963).

Content analysis is increasingly being used to assess psychological variables in the context of political decision making, particularly in the area of foreign policy. One approach, a continuation of the Lasswellian tradition, has emphasized elite values and ideology. In the absence of direct measures, Soviet and American publications representing political, economic, labor, military, scientific, and cultural elites were examined to identify major values. Themes regarding the economy, social and internal political affairs, and external relations were coded into more than 40 category sets. An investigation of elite foreign-policy attitudes for the same period focused on perceptions of the international system, power relationships, and operational codes. Although Soviet and American value preferences were found to be symmetrical in some respects, and incompatible in others, the data also revealed that elites in both nations displayed "a powerful tendency to act and speak in such a way as to exacerbate differences" (Angell, Dunham, and Singer, 1964, p. 473).

A second approach to foreign-policy studies has focused on analysis of documents written by officials holding key decision-making roles. The basic assumption is that foreign-policy decisions, like all decisions, are in part a product of the policy maker's perceptions; that if men define situations as real, they are real in their consequences. Again, the choice of content analysis is based largely on the inability to use more direct methods by which to assess the perceptions, attitudes and values of foreign-policy leaders at the time of decision; systematic analysis of diplomatic documents provides an indirect method to bridge gaps in time and space. Themes, which have been classified into categories such as friendship, hostility, frustration, and satisfaction, have served as the unit of analysis (North *et al.*, 1963).

An initial study tested two basic hypotheses about the relationship between perceptions of threat and perceptions of capability during an international crisis

(Zinnes, North, and Koch, 1961). During the weeks prior to war in 1914, perceptions of capability appeared much less frequently in decision makers' documents as perceptions of threat increased. This study also revealed the limitations, for many purposes, of using frequency as the sole basis of inference. After the 1914 data were recoded to permit analysis on the basis of intensity as well as frequency, hypotheses relating to perceptions of capability and injury were reexamined. Decision makers of each nation most strongly felt themselves to be victims of persecution and rejection precisely at the time when they were making policy decisions of the most crucial nature (Holsti and North, 1965).

Other analyses of the 1914 data, within the framework of a model linking actions and perceptions, have consistently shown that the more intense the interaction between parties, the more important it is to incorporate perceptual variables, as indexed by content data, into the analysis (Holsti, North, and Brody, 1968). This strong relationship between perceptions of hostility, feelings of involvement, and policy decisions has also been found in a study of Soviet and Chinese leaders during three crisis situations (Zaninovich, 1964).

A number of prominent hypotheses in the decision-making literature were tested using documents from the 1914 crisis. The data revealed that as stress increased, decision makers perceived time as an increasingly salient factor in formulating policy and they became preoccupied with the short-term, rather than long-range, implications of their actions. Leaders in various capitals of Europe also perceived the alternatives open to themselves to decrease, and those of their adversaries to increase, as they came under more intense stress (Holsti, 1965b).

In a study comparing events during the Cuban missile crisis of 1962 with those in the summer of 1914, some important differences emerged. During the Cuban crisis both sides tended to perceive rather accurately the nature of the adversary's actions, and then proceeded to respond at an "appropriate" level. Thus, unlike the situation in 1914, efforts by either party to delay or reverse the escalation were generally perceived as such, and responded to in a like manner (Holsti, Brody, and North, 1965).

To infer aspects of culture and cultural change

Anthropologists, sociologists, and others have traditionally examined societal artifacts to describe constant and changing characteristics of cultures. Content analysis for this purpose can be illustrated by a series of studies centering on hypotheses relating "need of achievement" and "inner/other direction" to major states in cultural development.

A person with high *n* Achievement is someone who wants to succeed, who is energetic and nonconforming, and who enjoys tasks which involve elements of risk; *n* Achievement has been defined operationally as "a sum of the number of instances of achievement 'ideas' or images" (McClelland, 1958, p. 520). The hypothesis that "a society with a relatively high percentage of individuals with high *n* Achievement should contain a strong entrepreneurial class which will tend to be active and successful particularly in business enterprises so that *the society*

TABLE 6
NUMBER OF *n* ACHIEVEMENT IMAGES PER 100 LINES BY TYPE OF SAMPLE BY TIME PERIOD

Period	Man and his gods	Estate management	Funeral celebrations	Poetry	Epi-grams	War speeches	Average
Growth 900-475 B.C.	2.01	3.54	7.93	2.87	4.72	7.38	4.74
Climax 475-362 B.C.	1.21	0.82	5.94	0.38	2.36	5.55	2.71
Decline 362-100 B.C.	0.81	0.00	2.54	0.16	1.57	3.00	1.35

will grow in power and influence" was tested by scoring samples of literature from the periods of growth (900-475 B.C.), climax (475-362 B.C.), and decline (362-100 B.C.) of Greek civilization (McClelland, 1958). As an index of economic power and influence, the location of vase remains was used to construct maps of the area within which Greece traded in the sixth, fifth, and fourth centuries B.C. These figures were 1.2, 3.4, and 1.9 million square miles, respectively. The amount of *n* Achievement imagery found in various kinds of literature is revealed in Table 6. When compared to trade area, the findings supported the hypothesis that expressions of *n* Achievement index stages in the development of a civilization

An independent check on these results was made by analyzing inscriptions on vases produced in various eras of Greek civilization. Aronson (1958) found that "doodling" styles can be used to discriminate persons with high *n* Achievement from those with low *n* Achievement. An objective scoring system for line shapes, and spaces of spontaneous "doodles" has been cross-validated against several groups of subjects. Without serious modification, the same scoring system was applied to inscriptions on Greek vases. The results substantiated the other findings; signs of high *n* Achievement were significantly more frequent in the period of growth and less frequent in the period of climax. The same system of content analysis has also been used in a cross-cultural study relating child-training practices to *n* Achievement in Indian folktales (McClelland and Friedman, 1958; for other examples, see McClelland, 1961).

According to some students of American culture, there has been a notable trend from the "Protestant ethic" or inner-direction to a "social ethic" or other direction (Riesman, Glazer, and Reuel, 1950). On the assumption that inner- or other-direction could be measured by achievement motive and affiliation motive (as defined by McClelland, 1958), respectively, children's readers for the period 1800-1950 were content-analyzed to determine whether these psychological variables index observed cultural change in the United States. Three hypotheses were tested: (1) that there has been a decrease in the incidence of achievement motivation and moral teachings; (2) that there has been an increase in occurrence of the affiliation motive; and (3) that incidence of the achievement motive has been positively correlated to an independent measure of achievement—the number

ber of patents issued. Achievement motivation in the readers increased steadily throughout the nineteenth century and began to decline only around the turn of the century; the same trend was found in the number of patents issued. The hypotheses relating to moral teachings and affiliation motivation were also supported (deCharms and Moeller, 1962).

A somewhat different test of the hypothesis of increasing other-direction was provided by a content analysis of advertising in a mass circulation women's magazine for the period 1890-1956 (Dornbusch and Hickman, 1959). The null hypothesis, that there was no difference in the proportion of other-directed advertisement before and after the midpoint of 1921, was rejected at the 0.001 level in eight tests.

These studies illustrate some of the many possible ways in which content analysis of social and historical documents can be used to test hypotheses. At the same time it should be pointed out that there are many pitfalls, aside from such technical problems as coding reliability, to be avoided. A most important problem, one rarely resolved beyond doubt, is the selection of materials which do in fact represent the culture, or at least some significant segment of it. Do newspapers, drama, or literature of a period, taken collectively, represent merely a manifestation of the authors' personalities, or do they reflect the more general milieu?

A partial solution to the problem is to rely on materials which meet the criterion of popularity, as was done in the study of achievement motivation in Greek literature. The rationale for this approach has been spelled out in a cross-cultural study of themes in popular drama (McGranahan and Wayne, 1948, p. 430):

Our first assumption in this study is that popular drama can be regarded as a case of "social fantasy"—that the psychological constellations in a dramatic work indicate sensitive areas in the personalities of those for whom the work has appeal; their needs, assumptions and values are expressed ("projected") in the drama. The successful play must be attuned to the audience.

A second approach involves examining materials which explicitly perform the function of transmitting and instilling social norms. Such materials, which may take very different forms across culture and time, have been widely used in content-analysis research; examples include folktales (McClelland and Friedman, 1952; Colby, 1966b), children's readers (deCharms and Moeller, 1962), youth manuals (Lewin, 1947), songs (Sebald, 1962), and textbooks (Walworth, 1938).

A third method is to use one or more independent indices against which to correlate content data. In their comparative study of German and American drama during the 1927 season, McGranahan and Wayne used six separate sets of data, both content and noncontent, to support their conclusion that there were real and persistent differences in the psychology of Germans and Americans (1948). A sample of plays from the 1909-1910 season was analyzed to determine whether differences could be attributed to Germany's defeat in World War I. Another test compared German audience reactions to movies which had been successful or unsuccessful in the United States. Each supplementary test supported findings based on content analysis, thereby increasing confidence in them.

To provide legal evidence

During World War II the United States government asked Harold Lasswell to analyze certain materials and to testify about their content in four cases of suspected criminal sedition. The purpose was to demonstrate that statements by the accused publishers conformed to enemy propaganda themes. Materials ranged from over two hundred books in English and Russian in the *Bookniga* case to eleven issues of the periodical *The Galilean*, published by William Dudley Pelley. Eight tests were developed to analyze the materials, the results of which were accepted in evidence by the court (Lasswell, 1949, pp. 177-178):

1. *Avowal: Explicit identification with one side of a controversy.*
2. *Parallel: The content of a given channel is compared with the content of a known propaganda channel. Content is classified according to themes.*
3. *Consistency: The consistency of a stream of communication with the declared propaganda aims of a party to a controversy. The aims may be official declarations or propaganda instructions.*
4. *Presentation: The balance of favorable and unfavorable treatment given to each symbol (and statement) in controversy.*
5. *Source: Relatively heavy reliance on one party to a controversy for material.*
6. *Concealed source: The use of one party to a controversy as a source, without disclosure.*
7. *Distinctiveness: The use of vocabulary peculiar to one side of a controversy.*
8. *Distortion: Persistent modification of statements on a common topic in a direction favorable to one side of a controversy. Statements may be omitted, added, or over- or under-emphasized.*

The most outspoken critics (St. George and Dennis, 1946; Hughes, 1950) of the content-analysis data presented in the sedition trials hardly qualify as dispassionate observers. Yet some of their specific objections are not wholly without merit. It has been pointed out that "There is almost no theory of language which predicts the specific words one will emit in the course of expressing the content of his thoughts" (Lasswell, Lerner, and Pool, 1952, p. 49). In the absence of such a theory, a posture of great skepticism is warranted toward use of content-analysis data for other than descriptive purposes in legal proceedings.

A form of content analysis, intended to yield only descriptive information, has been used by the Federal Communications Commission to determine whether radio-station owners conform to prescribed standards (Content Analysis . . . , 1948, p. 910). In its annual survey, the FCC compares station logs with ideal ratios between commercial and local live, sustaining, and public-issue programs. In one case the American Jewish Congress sought to deny an application by the *New York Daily News* for an FM broadcasting license on grounds of unfavorable bias against minority groups. In a split opinion the Commission ruled that both qualitative and quantitative data contained "technical deficiencies . . . so serious

as to vitiate any real value the analysis might otherwise have had." At the same time, the Commission ruled in unambiguous terms that content analysis is an acceptable evidentiary technique if the data are deemed to be of adequate quality (Content Analysis . . . , 1948, p. 914).

On the whole, content analysis has been used sparingly as a source of legal evidence. Literary infringement cases are perhaps the legal area in which it might be used most suitably. Existing tests suffer from precisely those deficiencies which can be remedied through careful content analysis (Sorenson and Sorenson, 1955, p. 264):

To distinguish the ideas, plots, title, phraseology, characters and locale, all of which are not infringing from the "original form of expression, language or thought sequence and literary style" is simply too difficult a job for the ordinary observer making a superficial comparison.

Meanwhile, copyright counsel have no way of knowing whether the judge read all of the works, scanned part of each, used his own sampling system, or how much weight, if any, he may give to the exhibits submitted.

The tests developed in the sedition cases, as well as those developed by "literary detectives," might well provide data better than impressionistic scanning.

THE ANTECEDENTS OF CONTENT: WHO

Who wrote *The Imitation of Christ*? Was James Madison or Alexander Hamilton the author of *The Federalist Papers* Nos. 49-58, 62, and 63? These are two of many problems of literary detection which have been investigated by content analysis. The belief that each person's style contains certain unique characteristics is an old one, and methods of inference from statistical description of content attributes go back at least to the nineteenth century (Mendenhall, 1887). But because there are so many possible characteristics of style which might be used to discriminate between authors, the major task is that of selecting proper indicators. The problem must often be tackled in a "fishing expedition" manner, as reliable discriminators in one case may fail in another. For example, sentence length, often a useful index, proved useless in the case of the *Federalist Papers*—the undisputed writings of Madison and Hamilton averaged 34.59 and 34.55 words per sentence, respectively.

Frequencies of various classes of nouns were used to determine whether Thomas à Kempis or Jean Gerson wrote *The Imitation of Christ*. Five independent tests, based on the incidence of approximately 8200 nouns each, yielded the results in Table 7 (Yule, 1944, p. 274). On this basis Yule concluded that Gerson was not the author. Although it is impossible to "prove" that *The Imitation of Christ* was written by Thomas à Kempis, these data are clearly consistent with the hypothesis of his authorship.

The frequency of 265 words in known writings of Madison and Hamilton served as the test for 12 *Federalist Papers* whose authorship was disputed (Motteler and Wallace, 1964). The data strongly supported the claim of Madison's authorship. The weakest odds in Madison's case were 80 to 1 on *Paper No. 55*;

TABLE 7

INCIDENCE OF SPECIAL NOUNS IN "THE IMITATION OF CHRIST" AND KNOWN WRITINGS OF THOMAS À KEMPIS AND JEAN GERSON

	<i>The Imitation of Christ</i>	<i>Thomas à Kempis</i>	<i>Gerson</i>
Test 1	671	709	912
Test 2	376	365	823
Test 3	59	58	162
Test 4	6	7	21
Test 5	0	1	24

No. 56 was next weakest, at odds of 800 to 1 for Madison. Politically important words turned out to be far less effective discriminators than the high-frequency "function words." This finding is consistent with one generalization which has emerged from other studies of the "unknown communicator" in painting, literature, and music; it is the "minor encoding habits," the apparently trivial details of style, which vary systematically within and between communicators' works (Paisley, 1964).

Other statistical methods for identifying authors from content characteristics have emerged from studies of the "Quintus Curtius Snodgrass Letters" (Brinegar 1963) and the "Junius Letters," a series of political pamphlets written in 1769-1772 (Ellegård, 1962).

The pitfalls of authorship studies can be illustrated by the controversy over the assertion that Paul wrote only five of the Epistles (Morton, 1963). The claim was based on an analysis of seven indicators assumed to be reliable as "finger prints of the mind" for discriminating style: *sentence length*, and the frequency of the *definite article*, *third-person pronouns*, the aggregate of all parts of the verb *to be*, and the words *and*, *but*, and *in*. Using these tests, Morton concluded that six authors wrote the 14 Pauline Epistles. However, subsequent tests have at least disproved the generality of the seven indicators; the same tests indicated that James Joyce's *Ulysses* was written by six authors, none of whom wrote *Portrait of the Artist as a Young Man* (Ellison, 1965). While not resolving the question of who wrote the Epistles, such results clearly raise doubts about Morton's findings.

THE RESULTS OF COMMUNICATION: WITH WHAT EFFECT

The basic format of content-analysis research designed to study the effects of communication is: If messages have attributes *A*, *B*, and *C*, then the prediction is that the effect on the recipient will be *x*, *y*, and *z*. Content analysis serves to describe the relevant attributes of the independent variables (*A*, *B*, and *C*). But as indicated earlier, any direct inference as to effects from content is at best tenuous. When a government-controlled newspaper such as the Chinese *People's*

Daily is analyzed to measure elite attitudes (Eto and Okabe, 1965), problems of inference are somewhat limited. The assumption that this source does reflect leadership views is plausible and can be verified by independent analyses of content or noncontent indices. But it would not necessarily follow that the same source could be used to measure mass opinion.

Evidence demonstrating that effects of communication are related not only to attributes of content but also to predispositions of the audience is too voluminous to review here (cf. Klapper, 1960; Bauer, 1964). A single example will suffice. Major themes of political appeals during the 1940 Presidential campaign were identified by content analysis of both public and private media. Reactions of the public to the content of the arguments were then measured by interview. The interaction between content and other factors is summarized by Berelson (1942, p. 63):

Why do people come across arguments and why do they accept them? Briefly, our answers are these. Mainly, people come across the arguments which the mediums of communication emphasize; they also tend to see the arguments they want to see and other arguments whose statement is appealing. Mainly (within a given time), people accept the arguments which support their own general position; they also tend to accept the arguments which they see in the public communications and those whose statement is persuasive.

Berelson's conclusion can be restated within the framework of content-analysis research: because of the variety of audience predispositions and decoding habits, the effects of communication cannot be inferred directly from the attributes of content ("what") or style ("how") without independent validation.

This problem was anticipated in an early proposal to measure public opinion by quantitative newspaper analysis. Woodward's (1934) research design incorporated systematic efforts to test the relationship between public attitudes and newspaper content. Often, however, this relationship is simply assumed to be a positive one. For example, frequencies of British and American place-names in colonial American newspapers were tabulated to index sentiments of national identity (Merritt, 1966). But the absence of evidence demonstrating that the appearance of these symbols either reflected or shaped public views calls into serious question inferences drawn from the content data. In short, the burden of proof is on the investigator to present evidence that audience attitudes can indeed be inferred directly from communications produced for that audience. As indicated in Table 1, measures of effects may be derived by analyses of (a) subsequent messages produced by the recipient to determine whether they are consistent with predicted effect, or (b) noncontent indices of the recipient's behavior.

To measure readability

Perhaps the most systematic content-analysis research measuring effects of communication has centered on correlating attributes of style with ease of comprehension. Characteristics of text which have often been tested include various aspects

of vocabulary (diversity, hard words, long words, abstract words), sentence structure (length, type, number of prepositional phrases or indeterminate clauses), and human-interest elements (personal pronouns, colorful words).

There have generally been two approaches to identifying and validating elements of readability. The first has been to identify the distinguishing attributes of materials prejudged to have a certain level of difficulty—for example, adventure stories and philosophical writings. This approach assumes that the investigator can judge *a priori* the level of difficulty of such materials—often a dubious premise. The more reliable technique has been to use independent tests to determine reading comprehension of subjects.

One of the earliest systematic studies on readability was undertaken by Gray and Leary (1935). They found that 24 of 44 language variables were significantly correlated with comprehension scores. After eliminating some variables which were highly correlated with others, five were selected to index readability: the number of different words, uncommon words, personal pronouns, prepositional phrases, and the average length of sentences.

Flesch (1943) and Lorge (1944) developed formulas based on three factors—sentence length, and two measures of vocabulary. The Flesch system was later revised, with formulas to measure both "reading ease" and "human interest" characteristics of text.

Subsequent research has focused in large part on developing easier methods of scoring the vocabulary factor in Flesch's formula. One method is to count only those words falling outside a standard 3000-word list (Dale and Chall, 1948). Another study demonstrated that incidence of one-syllable words is so highly correlated with the Flesch vocabulary factor that it may be used in its place, with a considerable saving of scoring time (Farr, Jenkins, and Patterson, 1951).

While readability research has consistently pointed to the importance of vocabulary load and sentence length (Brinton and Danielson, 1958), the limitations of these measures must be recognized. These formulas measure only *style* not other important elements of readability: content, organization, and format (Kearl, 1948). Second, readability measures cannot be applied mechanically; for example, it is true only to a point that the shorter the sentence, the more readable the text. Finally, these formulas are geared to the effect of style on the "general" audience; little allowance is made for the reader's experience and expectations each of which can have an important bearing on the extent to which the text is understood (Waples, Berelson, and Bradshaw, 1940, pp. 135-145).

The "Cloze Procedure," which represents a radical departure from traditional approaches to readability, overcomes the first two of these limitations. The text is initially mutilated by removal of every fifth, seventh, tenth, or *n*th word after which the reader is asked to supply the missing words (Taylor, 1953). The index of readability is based on the percentage of blanks correctly filled in. The advantage of the method is that it is effective in cases where other methods break down; that is, when idiosyncratic use of language (for example, the writings of James Joyce or Gertrude Stein) produces invalid readability scores. Although

originally developed to measure readability, because the Cloze system is sensitive to semantic, associational, grammatical, and syntactical determinants of verbal behavior, it has subsequently proved useful in personality (Honigfeld, Platz, and Gillis, 1964) and psychiatric research (Fillenbaum and Jones, 1962; Salzinger, Portnoy, and Feldman, 1964).

Other attributes of style whose effects have been studied, and for which measures have been developed, include "sensationalism" (Tannenbaum and Lynch, 1960) and "abstraction" (Gillie, 1957; Haskins, 1960).

To analyze the flow of information

The flow of news to the United States from the outbreak of World War I to America's entry in April 1917 was analyzed by Foster (1935, 1937). Over 11,000 items appearing on the front page of the *New York Times* and in the Chicago press were coded according to origin of the news, and to type of appeal contained within it which might make the reader favor American participation. The data revealed that American readers were almost wholly dependent on news directly from, or dispatched through, the Entente powers. Thus, events such as the German invasion of Belgium were reported almost exclusively by news received from Germany's enemies. As war approached, the proportion of news from American sources increased sharply, as did news containing some appeal favoring American participation.

Content analysis was used to determine whether the Associated Press and United Press International, each of which had full-time bureaus in Havana, were responsible for charges of inadequate public information about the Cuban revolution. All stories about events in Cuba filed during December 1958 were analyzed. Tables summarizing AP and UPI reports were compared to coverage of the Castro revolution by major newspapers published in Washington, Cleveland, and Louisville. Scores were computed for percentage of available AP and UPI information used, and the prominence (headlines, placement) with which it was displayed. On the basis of the comparison, the author absolved the news services from charges of inadequate coverage: "The newspapers received enough wire copy to tell the long, continuing story of the Cuban revolution. They made little use of this material, however, until the last six days" (Lewis, 1960, p. 646).

A form of content analysis was used to study the transmission of rumors through six removes from an original source (De Fleur, 1962). Subjects in a community were informed that prizes would be offered to those who could repeat a short advertising slogan to a team of investigators. The slogan was given to a sample of subjects, and interview responses three days later were analyzed to determine the degree of distortion in the original message. By asking subjects to identify their source of information, it was possible to determine how many steps removed each respondent was from the original source.

Studies employing some form of content analysis to chart flow of news include Carter (1957), Schramm (1959), Hart (1961), and Galtung and Ruge (1965).

To assess responses to communication

One aspect of the effects of a communication is the degree to which its symbols become assimilated by the audience. Prothro tested the hypothesis that political symbols of the New Deal have become a "permanent increment to the main body of the American tradition," and that not even successful spokesmen for conservatism reject them (1956, p. 727). The first Acceptance, Inaugural, and State of the Union addresses of Presidents Hoover, Roosevelt, Truman, and Eisenhower were coded for relative frequencies of "political appeals" (government aid, government regulation, national power, etc.) and "demand symbols" (peace, freedom, faith, controls, initiative, etc.). While demand symbols distinguished Hoover and Eisenhower from Roosevelt and Truman, Eisenhower's political appeals were free from any repudiation of the New Deal, thereby supporting the hypothesis.

Soviet newspapers and domestic and foreign broadcasts were analyzed to assess the effects of Voice of America broadcasts. For a four-year period beginning in 1947, mass communication materials were examined for all references to the Voice of America. A number of content-analysis techniques were used to answer different questions. Frequency counts were used to measure focus of attention and distribution of Soviet references to VOA. Most foreign attacks on VOA were directed to Eastern and Western Europe, with little attention to Latin American audiences. During the four-year period there was a relative increase in attention directed to domestic audiences, especially in those publications read predominantly by the Soviet intelligentsia. Thematic analysis was used to code more than 2500 references to VOA. These data revealed that the Soviets, rather than posing a counterimage of Soviet virtues, responded by counteracting the image they assumed VOA had created of the United States (Inkeles, 1952; see also Massing 1963).

Does the content of communication have a greater effect on the audience when it is attributed to a high-prestige source? Attitudes before and after exposure to three messages attributed to Thomas Gates, Walter Lippmann, and James Conant indicated that messages with a by-line do produce greater change in the direction of the message, but only when the by-line is located at or near the top of the message (Tannenbaum and Greenberg, 1961a; see also Hovland, Janis, and Kelley, 1953).

Two studies have examined the effects of scholarship by analyzing published reactions to research. Lerner undertook a thematic content analysis of the published reviews of *The American Soldier*, a study of Army morale during World War II by a large team of social scientists. Responses were related to professional affiliation and other attributes of the reviewer. In his conclusion, Lerner suggested how content and predispositions combined to shape the nature and tone of reviews (1950, pp. 241-242):

By the test of response, *The Soldier* surely is an important book. The challenge it presents to the stock attitudes of readers is clear from the depth of the affects exhibited

by its reviewers: the responsive postures range from piety to diabolism. . . . The lack of time, skill, or inclination to refigure one's position in *The Soldier's* universe of discourse meant that most reviewers had to draw upon their already available stock of attitudes. This leads to "spontaneous" self-positioning, which accounts for both the regularity of alignment by occupation (i.e., prior commitment as to scientific research in human affairs) and the readiness to displace affective responses from the concrete object *The Soldier* to the general symbol Social Science.

A somewhat similar study examined the treatment of Freud's theory of dreams in general psychology, abnormal psychology, and psychiatry textbooks (Herma, Kris, and Shor, 1943). The investigators initially developed a list of 30 propositions basic to the theory of dreams. All available textbooks published between 1901 and 1940 were coded according to: mention of the theory of dreams; frequency of basic propositions cited; attitude toward the theory as a whole and toward specific propositions comprising the theory; and, if the theory was rejected, the basis of rejection (ridicule, moral grounds, scientific grounds, etc.). The data revealed that the theory of dreams had become a predominant interest only in texts on abnormal behavior, and that those who rejected the theory did so on grounds other than scientific.

CODING CONTENT DATA

Coding is the process whereby raw data are systematically transformed and aggregated into units which permit precise description of relevant content characteristics. Coding rules serve as the operational link between the investigator's data and his theory and hypotheses. Coding rules are thus an integral part of the research design, and in preparing them the analyst makes a number of decisions. Specifically:

- How is the research problem defined in terms of categories?*
- What unit of content is to be classified?*
- What system of enumeration will be used?*

Although coding is discussed under these three headings, the division is solely for purposes of exposition. It does not imply that selection of categories, units, and system of enumeration are independent decisions. These represent a series of interrelated choices, each of which carries with it certain assumptions.

CATEGORIES OF ANALYSIS

A central problem in any research design is selection and definition of categories, the "pigeonholes" into which content units are to be classified: "Content analysis stands or falls by its categories. Particular studies have been productive to the extent that the categories were clearly formulated and well adapted to the problem and to the content" (Berelson, 1952, p. 147). There are as many possible schemes for classifying content data as there are questions which may be asked of the data. Among the types of categories used frequently in content-analysis

research are the following (some of which are described in greater detail in Berelson, 1952, pp. 147-168):

"What is said" categories:

- Subject matter.* What is the communication about?
- Direction.* How is the subject matter treated (for example, favorable-unfavorable strong-weak)?
- Standard.* What is the basis on which the classification by direction is made?
- Values.* What values, goals, or wants are revealed?
- Methods.* What means are used to achieve goals?
- Traits.* What are the characteristics used in description of people?
- Actor.* Who is represented as undertaking certain acts?
- Authority.* In whose name are statements made?
- Origin.* Where does the communication originate?
- Target.* To what persons or groups is the communication directed?
- Location.* Where does the action take place?
- Conflict.* What are the sources and levels of conflict?
- Endings.* Are conflicts resolved happily, ambiguously, or tragically?
- Time.* When does the action take place?

"How it is said" categories:

- Form or type of communication.* What is the medium of communication (news paper, radio, television, speech, etc.)?
- Form of statement.* What is the grammatical or syntactical form of the communication?
- Device.* What is the rhetorical or propagandistic method used?

This list is neither exhaustive nor does it define the limits of content analysis. It is merely an enumeration of categories which have been employed more or less frequently. Nor are the categories at the same conceptual level. Subject matter categories can be used independently, as is done when newspaper content is classified into various types of news. But one cannot code items for direction (for example, approval-disapproval) independently of a referent. The same is true of the "standard" category. The coder may first be asked to isolate a subject matter unit (for example, Freud's dream theory), then classify it for direction (attitude (favorable or unfavorable), after which some judgment is made as to the standard used for acceptance or rejection (Herma, Kris, and Shor, 1943).

One of the questions frequently raised in the literature is that of standard categories. The advantages of standardization are the same as in any area of scholarship: results may be compared across studies and findings will tend to become cumulative. On the other hand, the disparity of purpose which characterizes content-analysis research makes standardization difficult to achieve. Some categories have been rather widely employed in descriptive studies of newspaper content (Bush, 1961; Willey, 1926; Woodward, 1930), values (Lasswell, 1930; White, 1951), political symbols (Pool, 1952a), attitudes (Osgood, Suci, and Tannenbaum, 1957), and a few other areas. There have also been recent attempts

to develop standard categories for general psycholinguistic analysis (Laffal, 1965) and dream analysis (Hall and Van de Castle, 1966). But in general, Pool's observation (1959, p. 213) accurately reflects the current state of the field:

It is questionable, however, how ready we are to establish standard measures . . . in content analysis. Such a measure is convenient when a considerable number of researchers are working on the same variable, and when someone succeeds in working out good categories for that variable. It is doubtful that either of those criteria can be met in most areas of content analysis.

This state of affairs, understandable as it may be, has effectively prevented development of content norms for most classes of communicators. The absence of norms, in turn, often presents the investigator with knotty problems of inference. Do his findings regarding the communication content of a single source represent a significant deviation from those of the general class of communicators (for example, metropolitan dailies, mass circulation magazines, politicians, schizophrenics, etc.)? In view of the importance of such questions, formulation of standard categories and content norms would appear to be a high-priority area for future investigations.

In the absence of standard schemes of classification the analyst often is faced with the task of constructing appropriate categories by trial-and-error methods. This process usually consists of moving back and forth from theory to data, testing the usefulness of tentative categories, and then modifying them in light of the data. It is important to bear in mind, however, that the choice of categories can significantly affect one's findings. In a content analysis of American colonial newspapers a frequency count revealed an increase in American place-names after 1763, with a concomitant decline in the appearance of British symbols. But when references were classified as favorable, unfavorable, or neutral, changes were virtually nonexistent, even during a thirty-year period (Merritt, 1966, p. 49).

Definition of categories requires that they actually represent the elements of the investigator's theory; that they be exhaustive, to ensure that every item relevant to the study can be classified; and that they be mutually exclusive, so that no item can be scored more than once within a category set. To ensure results which are replicable, the investigator must specify explicitly the *indicators* that determine which units fall into each category. How, for example, is the coder to recognize a statement of "need achievement" (McClelland, 1958) or indications of anti-intellectualism in political reporting (Hage, 1959)? Rarely are categories so self-evident that coders will not require clear and specific instructions to guide their judgment.

A category can be defined exhaustively by enumerating each content unit to be placed in the category. In the RADIR studies the coders were supplied with a list of 416 key symbols which defined the limits of the required information (Pool, 1952a). The coding process was thus reduced from a judgmental task to a clerical one. Most categories, however, do not lend themselves to exhaustive definition, especially if units more inclusive than the word are used. Each category must then be defined more or less precisely by characterizing its

major properties. These are the rules which the coder uses to determine whether units fall within the boundaries of the category.

UNITS OF ANALYSIS

Recording units

In addition to defining the categories into which content data may be classified the analyst must designate the size of the units to be coded. The initial decision is the choice of *recording unit*, the specific segment of content that is characterized by placing it in a given category.

A number of recording units have been used frequently in content-analysis research.

The single word or symbol. In the past this unit has often been avoided in mass media research involving a great volume of data, a notable exception being the RADIR studies, in which nearly 20,000 editorials were analyzed. More recently computer content-analysis programs have materially reduced the costs and increased the reliability of analyses based on word frequencies. The word has more frequently been used in research on readability (Gray and Leary, 1935; Flesch, 1948; Taylor, 1953), style (Skinner, 1939; Miles, 1951; Parrish, 1959), psychotherapy (Dollard and Mowrer, 1947), and literary detection (Yule, 1944; Mostell and Wallace, 1964).

The theme. For many purposes the theme, a single assertion about some subject, is the most useful unit of content analysis. It is almost indispensable in the study of propaganda, values, attitudes, and the like. A major drawback to coding themes is that it is usually time-consuming. Another difficulty is that it is not easily identified "natural" unit like the word, paragraph, or item; thus the coder must make an intermediate judgment to identify the boundaries of the theme. The sentence, "These clandestine Soviet actions on the imprisoned island of Cuba will not be tolerated by the American people," contains assertions about three nations. The coder must be able to reduce this sentence into its component themes before they may be placed in the proper categories.

The character. Studies of entertainment materials in the mass media have often employed the character as the recording unit. In this case the coder tallies the number of persons, rather than the number of words or themes, into the appropriate categories. Such research has focused on ethnic, socioeconomic, marital, psychological, and other traits of characters as portrayed in magazines (Berels and Salter, 1946), movies (Jones, 1942), television (De Fleur, 1964), comics (Spiegelman, Terwilliger, and Fearing, 1953a), and other products of the mass media. Similar units have been used to examine traits of characters appearing in novels or drama for the purpose of drawing inferences about the personalities of authors (Kanzer, 1948; McCurdy, 1939, 1947, 1953).

The paragraph, the sentence, or other grammatical units. In part because these units often do not lend themselves to classification in a single category, they have

rarely been used. The sentence about Soviet actions in Cuba, cited above, illustrates a problem that is even more severe when the paragraph is the recording unit.

The item. In this case the entire article, film, book, or radio program is characterized. This unit is too gross for most research, and may present problems when items fall between two categories; for example, is a war film with a comic theme classified under "war" or "comedy?" Item analysis is most useful for coding great amounts of materials when gross categories will suffice.

Context units

It may not be possible to classify a recording unit without some further reference to the context in which it appears. Attitudes toward democracy cannot be inferred solely on the basis of how frequently the word itself—and others defining the category "democracy"—appears in the communication; the *context unit* is the largest body of content that may be searched to characterize a recording unit. The coder may be instructed to refer to the sentence, the paragraph, or the entire document, in order to determine the attitude toward the symbol democracy.

Selection of recording and context units should rest upon two considerations. First, which units will best meet the requirements of the research problem? This question can only be answered in light of the hypotheses and the nature of the data. The important point is that *the units chosen may affect the results of the analysis.* Geller, Kaplan, and Lasswell (1942b) performed an experiment comparing four different recording and context units on the same sample of editorial matter.

Method	Recording unit	Context unit
I	Symbol	Sentence
II	Paragraph	Paragraph
III	3 sentences	3 sentences
IV	Article	Article

The four methods were in agreement in indicating *direction* of bias (favorable, unfavorable, neutral), but differed considerably in revealing its *extent*. In general, the larger the recording unit, the more the degree of bias in editorials was emphasized, and as the size of context unit was increased, the number of neutral entries diminished.

The effect on one's findings which may result from the choice of context units can be further illustrated in *contingency analysis*. In this method of content analysis, inferences are based on co-occurrence of attributes within the same unit. If the context unit is very small, few co-occurrences will be found; on the other hand, if the entire message is the context unit, everything which appears will be contingent with everything else. Osgood very tentatively suggested that stable results can be achieved with units of 120 to 210 words (Osgood, 1959, p. 62).

A second factor which few investigators can afford to overlook is efficiency; that is, which units give satisfactory results with the least expenditure of resources?

In an analysis of best-selling inspirational literature, themes appearing in each paragraph were initially coded, a task which proved disappointingly time-consuming and difficult (Schneider and Dornbusch, 1958, Appendix C). A second coding method was also used: the entire book was read and assigned a single summary score for each category. A comparison of results derived by the two methods revealed that little substantive information was lost with the latter approach. Questions of efficiency can often be answered only by comparing two or more methods on the same set of data. Unfortunately, relatively few studies have reported experiments of this kind.

SYSTEMS OF ENUMERATION

The analyst must also choose the unit of enumeration, that is, the unit in terms of which quantification is to be performed. The recording unit and the unit of enumeration may be identical, as was the case in a study of occupations portrayed on television; the occupation of each character was recorded and tallied (De Fleur 1964). In a similar study of soap operas, recording and enumeration units were different. The occupation of each character was recorded, but the results were reported according to the number of scenes in which any occupational type appeared; thus a scene involving three housewives was tallied only once (Arnheim, 1944).

Systems of enumeration vary considerably in precision and in time required to code a given sample of data. The investigator must determine how fine are the discriminations he needs in order to satisfy the requirements of his problem generally, the greater the need for precision, the higher will be the costs of the analysis. Often the nature of the categories and data are such that the search for maximum precision will not only entail considerably higher costs, but also may sacrifice reliability. The choice may also affect the results of the analysis. For example, editorials may be coded according to the frequency of favorable or unfavorable assertions about a specific issue. If the unit of enumeration is the single assertion, results will indicate whether the editorial was only slightly (50+%) or predominantly on one side or the other. If the entire editorial serves as the unit of enumeration, that is, if the whole editorial is scored pro or con this distinction will be lost.

The most important aspect of the investigator's choice, however, is that each system of measurement carries with it a certain set of assumptions regarding the nature of the data and inferences which may be drawn therefrom.

Time/space

Most early studies employed measures of space (for example, column inches) to describe relative emphases in the content of newspapers. The analogous unit for film, radio, and television is time. The popularity of space/time measures can be attributed largely to the relative ease and reliability with which they may be used. For some purposes space/time measures may serve as an adequate substitute for other more time-consuming methods. Markham and Stempel (1957

compared newspaper political coverage as measured by the amount of space in column inches, frequency of issues mentioned, and size of headlines. The three methods yielded similar results, but coding for space was done considerably more rapidly than coding of issues. The limitations of space/time units derive mostly from a lack of sensitivity to other than the grossest attributes of content. In general, such measures are most appropriate for descriptions of mass media, but are too imprecise to index attitudes, values, style, and the like. For example, the use of space/time measures to answer questions about the degree of political bias in newspapers is at best questionable, inasmuch as a one-to-one relationship between focus of attention and bias cannot be assumed.

Appearance

One alternative to space/time units is that of searching the document, or some subunit of the text, for appearance of the attribute. The size of the context unit determines the frequency with which repeated items occurring in close proximity to each other are counted separately. Depending on the context unit, repetition of a given attribute within a sentence (White, 1947), paragraph (Schneider and Dornbusch, 1958), or item (Lasswell, Lerner, and Pool, 1952) does not change the tally. This method of enumeration has two important advantages. It can usually be done with relative ease and with high reliability because the coder is faced with a dichotomous decision (appearance-nonappearance). Moreover, this method is useful if one cannot assume a linear relationship between frequency and importance of content attributes. Some investigators have labeled this type of nominal measurement as "qualitative" content analysis, although the term is somewhat misleading because data coded in this manner may be presented numerically (for example, the percentage of items in which a given theme appears), as well as subjected to certain statistical tests.

Frequency

The most commonly used method of measuring characteristics of content is that of frequency, in which every occurrence of a given attribute is tallied. For example, how frequently do the folktales of various Indian tribes express need to achieve (McClelland and Friedman, 1952), or, how frequently did Madison and Hamilton use the words "whilst" or "upon" (Mosteller and Wallace, 1964)?

Again, the important point is that the analyst using a measure of frequency to test his hypotheses incorporates two related assumptions into his research design. First, he assumes that the frequency with which an attribute appears in messages is a valid indicator of some variable such as focus of attention, intensity, value, importance, and so on. Second, he assumes that each unit of content—word, theme, character, or item—should be given equal weight with every other unit, permitting aggregation or direct comparison.

George (1959b) and others have contended that, for purposes of drawing inferences about causes and effects of communication, the first premise is often untenable. Other questions have been raised regarding frequency measures.

For example, what is the relationship between distribution, or permeability, attributes in the text and their importance (Dunphy, 1964)? Is there a difference between two variables appearing equally often if their distribution in the documents under analysis is different?

The dubious validity, for many research problems, of the assumption that each unit should be assigned equal weight has more often received recognition from content analysts (cf. Stewart, 1943). Descriptions of newspaper content, for example, routinely differentiate items appearing on the front page from those printed elsewhere, either by coding them separately, or by using a system of "weighting" to reflect the factor of prominence. Similar devices have been used to reflect the position of items within the page and size of headlines. Ash's (194) study of the Taft-Hartley Act illustrates a method of adjusting scores to reflect circulation. An evaluation of the act received a score proportional to the circulation of the magazine in which it appeared.

Intensity

For research dealing with values and attitudes, a serious problem engendered by the use of pure frequency counts is that of intensity. In other words, the assumption that valid inferences regarding attitudes can be drawn from frequency scores, unadjusted for intensity of expression, is often untenable. Some of the problems can be illustrated by examining four Soviet statements:

The Chinese may be preparing to denounce Khrushchev's policies.

The Chinese sometimes disagree with Khrushchev's policies.

The Chinese will soon begin denouncing Khrushchev's policies.

The Chinese are bitterly denouncing Khrushchev's policies.

An analyst wanting to map trends in Sino-Soviet relations would require some method of differentiating the intensity of the four themes.

Categorization of content units for intensity normally presents more problems than other coding processes. Construction of scales is usually a difficult process. Moreover, the broad range of linguistic elements which indicate intensity makes it difficult to list all criteria the coder may have to consider in making his decisions. In the themes cited above, there are at least four elements to be considered in judging intensity: (1) the relative intensities of the verbs "disagree" and "denounce"; (2) the function of the adverbial modifiers "sometimes" and "bitterly"; (3) the element of tense—past, present, and future actions; (4) the probabilistic character of the first theme. This list by no means exhausts the intricacies of language which may denote intensity.

One method of coding for intensity is the "paired comparison" technique developed by Thurstone. Judges decide which of each possible pair of indicators is rated higher on a linear scale of attitudes. The judgments are used to construct categories into which content units are placed. A simplified version of this procedure was used to study national attitudes, as revealed in newspapers and journals, during two crisis periods in the Far East (Russell and Wright, 193

TABLE 8
EVALUATIVE ASSERTION ANALYSIS: AN ILLUSTRATIVE EXAMPLE

Soviet rulers are ruthless, atheistic despots. These men have in the past pursued evil goals. Yet there now appears some possibility that they will agree to some measures designed to relax world tensions. Perhaps they will be more willing to forgo aggressive designs.

Attitude object	Verbal connector	Value	Common-meaning term	Value	Product
1. Soviet rulers	are	+3	ruthless	-3	-9
2. Soviet rulers	are	+3	atheistic	-3	-9
3. Soviet rulers	are	+3	despots	-3	-9
4. Soviet rulers	have in the past pursued	+2	evil goals	-3	-6
5. Soviet rulers	may now possibly agree to	+1	some measures designed to relax world tensions	+2	+2
6. Soviet rulers	perhaps will be more willing to forego	-1	aggressive designs	-3	+3

Wright and Nelson, 1939). One study has demonstrated that scales derived empirically, as by the Thurstone technique, sometimes yield more satisfactory results than logical scales when relatively untrained coders are used. Differences between the two methods of scale construction were not significant when experienced and trained personnel did the coding (Exline and Long, 1965).

The Q-sort scaling technique has also been applied to content data. Judges are instructed to place content units into a fixed-distribution nine-point scale (North *et al.*, 1963). This is essentially a rank-order method, the rank of any unit relative to all other items determining its intensity score. Both paired-comparison and rank-order methods depend on the assumption that content units are sufficiently homogeneous on a single continuum that they may usefully be compared.

An "atomic" approach to coding for intensity, "Evaluative Assertion Analysis," has been developed by Osgood and his associates (Osgood, Saporta, and Nunnally, 1956; Osgood, 1959). The initial step is to translate all sentences into

one of two common sentence structures:

Attitude object / *Verbal connector* / *Common-meaning term*
*Attitude object*₁ / *Verbal connector* / *Attitude object*₂

For example, the sentence, "An aggressive Soviet Union threatens the United States," is translated to read:

The Soviet Union / *is* / *aggressive*
The Soviet Union / *threatens* / *United States*

Attitudes are computed on the basis of values assigned to verbal connectors and common-meaning terms. These range from +3 to -3 depending on their direction and intensity. The method is illustrated more fully in Table 8.

The final step is the computation of values for each attitude object, in this case, a single one—Soviet rulers. Extensive rules for coding and scoring have been formulated (Osgood, Saporta, and Nunnally, 1956).

Because this method first reduces the theme to its parts and specifies which elements of the theme are to be scored, it can be used with a high degree of reliability. Coders can be trained rapidly, but the method is too laborious to be used for large volumes of data, and it is uneconomical if only gross measures of attitude (for example, pro or con) are required. It is probably most useful when the analyst requires precise data on only a limited number of attitude objects; it has been used, for example, to assess the treatment of Presidential candidates by three major news magazines during the 1960 campaign (Westley *et al.*, 1963), to examine John Foster Dulles' attitude toward the Soviet Union (Holsti, 1967) and to analyze editorial treatment of India in the *New York Times* (Lynch and Effendi, 1964).

Other methods of accounting for intensity in content data are described by Jacob (1942) and Kaplan and Goldsen (1949).

SAMPLING, RELIABILITY, AND VALIDITY

The purpose of content-analysis research is to present a systematic and objective description of the attributes of communication. These data may be used to make inferences about communicators or audiences. But whatever the specific purpose of the study, there are certain problems which the content analyst shares with all social scientists. What is the universe of communication to be described, and what sample is to be drawn therefrom? Do independent measures of the same data yield results capable of verification within stated confidence limits? Do the categories actually index the variable they are intended to measure? Careful attention to these questions must be an integral part of every research design if it is to meet meaningful standards of systematic investigation. A comprehensive discussion of sampling, reliability, and validity is beyond the scope of this chapter; the present review considers only some general issues of particular relevance to content analysis.

SAMPLING

Because communication pervades society, the investigator is always faced with a problem of selecting some portion of it for his research. Content-analysis findings are usually discussed, implicitly or explicitly, as being relevant for some universe beyond the specific documents under study. Hence, both practical requirements of narrowing data to manageable proportions and problems of generalization present the analyst with sampling decisions.

Choices at any one step in the research design are not independent of those made at earlier stages. Thus, once the research problem has been defined, the sampling design has been partly determined. The investigator comparing coverage of a Presidential election in the prestige and mass press has already limited his data according to time period and class of communicators.

Ideally, the next step should be to list all members of the universe of communication from which a sample is to be drawn, in order to minimize the probability of systematic sampling error. But this is not possible in every case. The analyst of American daily newspapers has a relatively simple task; he may, for example, define his universe as those newspapers catalogued in *Editor and Publisher Yearbook*. The psychologist working with personal documents is faced with a considerably more difficult problem. In a small-group experiment the investigator has access to the entire universe of the subjects' communication in that situation, from which a true probability sample can be drawn. On the other hand, in studies of political communication during an international crisis one rarely has access to more than the written messages; oral communication represents an inaccessible, and generally unknown, proportion of the whole. Moreover, misfiling or destruction of documents, and bias or carelessness by those commissioned to collect and publish the documents, may further complicate matters.

Once the universe of relevant communication has been defined, a single-stage sampling design may suffice. More often, a multistage sample is required. This may involve as many as three decisions: selecting *sources* of communication, sampling *documents*, and sampling *within documents*.

Selecting communication sources

The first sampling decision is that of sources. Which newspapers, magazines, books, authors, speeches, broadcasts, movies, etc., will be selected as representative of the universe? The analyst may draw a random sample by one of several standard methods. This procedure is applicable when every source can be considered equally important for purposes of the study. If this assumption is not warranted, a purposive sample may be used to reflect qualitative or quantitative aspects of the sources which are deemed important. Studies of newspaper content have focused on prestige papers and journals because they were felt to represent most adequately the views of political elites (Lasswell, Lerner, and Pool, 1952; Angell, Dunham, and Singer, 1964). In studies of decision makers' attitudes

during international crises, communicators were selected on the basis of role; thus heads of state, heads of government, and foreign ministers were selected, but ambassadors were excluded (Holsti, North, and Brody, 1968; Holsti, Brody, and North, 1965). Purposive samples have also been selected by relying on expert opinion. In studies of psychological research (Allport and Bruner, 1949) and of newspaper coverage of the 1960 election (Stempel, 1961), sources rated most important by professional psychologists and journalists, respectively, were analyzed.

However, for other purposes quantitative criteria are more important than qualitative ones. Kayser (1953) analyzed the largest-circulation newspaper in each of 15 countries; Schneider and Dornbusch (1958) examined only that sample of inspirational literature on best-seller lists; and McGranahan and Wayne (1949) confined their study of German drama to the most widely attended plays. Other criteria which have been used to select a sample of communication sources include editorial position (Hage, 1959), geographical location (Willey, 1926; Tauber, 1932), and time of issue (Danielson and Adams, 1961).

Stratified sampling is often used in studies of the mass media to permit weighting of units according to certain criteria. A technique which permits stratification on the basis of circulation and geographical location has been developed (Maccoby, Sabghir, and Cushing, 1950). All dailies in the United States were originally listed in descending order of circulation within each of the nine census districts. The sampling rate was obtained by dividing the number of newspapers desired for the study into the total newspaper circulation, providing "circulation units." Each newspaper in which the n th circulation unit appeared was chosen—with the first paper selected from a table of random numbers—yield a true probability sample of circulation units in the universe. Another successful stratified sampling design for newspapers is described in Coats and Mulkey (1950).

Sampling documents

After communication sources have been selected, the investigator may have reduced his data to manageable proportions. If the universe of communication is still too large, the next step is the selection of documents. Here "document" refers to every item comprising the universe of items produced by the communication source. For example, if every Nazi broadcast to England during World War II is the universe, the investigator may decide to examine only a subset consisting of every third broadcast. In preparing a sampling design the analyst must weigh a number of factors. First, how large should the sample be to permit generalization within specified confidence limits? The necessary sample size varies depending on the kinds of questions that are being asked of the data and on the nature of the data. Though there is no answer to the question of sample size which can be applied to every problem, some guidelines are available from experience. Four methods of sampling newspaper headlines in *Pravda* were compared with

data for the entire month (Mintz, 1949). The results revealed that every-fifth-day samples (5th, 10th, 15th, 20th, 25th, and 30th day of each month) and odd-day samples did not differ significantly from the figures for the entire month. On the other hand, weekly samples and every-tenth-day samples (5th, 15th, and 25th day of each month) were inferior. These findings were supported by another study, in which every-sixth-day samples provided sufficiently accurate results for research purposes (Davis and Turner, 1951). Stempel (1952) drew samples of 6, 12, 18, 24, and 48 issues of a newspaper and compared the average content for a single subject matter category against the average for the entire year. The data indicated that each of the five sample sizes was adequate, and that increasing the sample size beyond 12 did not produce significantly more accurate results.

The analyst must also ensure that his sample, whatever its size, is free of any regular cycles within the universe which may bias results. Some types of content may be subject to seasonal variations; news about Congress will be heavier during the first half of the year than during the second, but the reverse is usually true about news of the United Nations General Assembly. Unless taken into account in the sampling design, such variations can render invalid an otherwise well-designed and executed project. A study of marriage announcements in the *New York Times*, based on Sunday issues of the newspaper appearing in June during the years 1932-1942, revealed no announcement of a marriage in a Jewish synagogue (Hatch and Hatch, 1947). A critic later pointed out that June almost invariably falls within a period during which tradition prohibits Jewish marriages (Cahnman, 1948). There are also regular variations within the week which may affect the content of mass media; for example, a study of Monday newspapers for financial news or news regarding governmental agencies would result in serious underestimation of coverage.

Several methods of accounting for variations within the week can be used. Every-*n*-th-day samples will ensure an equal distribution of days within the week if two conditions are met: *n* must not equal seven; and the number of weeks studied must be divisible by *n*. Jones and Carter (1959) have developed a "constructed week" procedure in which the period under study is divided into seven subpopulations; all Mondays are grouped together, as are all Tuesdays, and so on. A random sample of *n* newspapers is then taken from each.

A final consideration is that of efficiency. Which method of sampling will yield satisfactory results at least cost? Again, no definite answer covering every situation is possible. A case in point is the "cluster" design, in which the analyst selects units containing more than one item: the unit identifies a cluster consisting of all items appearing in it. For example, rather than listing all *articles* and drawing a sample from them, the investigator may sample *issues* of the magazine, and then analyze all articles which appear within the selected issues. Economy results from the reduced cost of listing. However, lower costs must be weighed against other factors. An experiment with content data revealed that, depending on the data, cluster sampling frequently results in overestimating levels of significance (Backman, 1956).

Sampling within documents

Even after selecting documents, the investigator may wish to reduce his data further by some sort of sampling *within* the document; for example, he may restrict his study to 30 pages drawn at random from a book (Harvey, 1953), to only the front page of a newspaper (Kingsbury, Hart *et al.*, 1937), or to every other story in a magazine (Berelson and Salter, 1946). The problem is the same as for other sampling decisions: does the selected sample accurately represent the attributes of content relevant to the study? The front page of newspapers may be a valid sample for some purposes, but it would not accurately reflect the volume of advertising or sports news.

In conclusion, it is worth repeating that sampling procedures applicable to all projects cannot be specified; nearly every content-analysis study presents a different problem. Furthermore, the choice between sampling designs—each of which carries with it certain assumptions—may significantly affect the results (Kaplan, 1964, p. 244):

What is important is the recognition that each sampling plan calls for its own standardization, and its own way of computing stability. As with measurement where there are many different scales that might be used, each having its own properties, sampling is not a simple uniform procedure but one which varies from problem to problem in a way that permits and even demands correspondingly different mathematical treatment.

RELIABILITY

If content analysis is to meet the requirement of objectivity, results must be reliable; that is, the research must yield results capable of verification by independent observers. The degree to which a given study will prove reliable is a function of the judges' skill, insight, and experience, and the categories into which content data are to be classified. The content analyst is thus concerned with the reliability of both coders and categories, each of which is important to the overall results of the research.

Individual reliability

Individual reliability reflects the extent of agreement between any coder and the rest of the judges. Before the actual coding begins, the investigator may want to run experiments to identify and eliminate judges deviating consistently from the group. This can be done by tabulating the correlation or percentage of agreement between every pair of judges. Even assuming that judges possess the skills necessary to make the discriminations required in the coding process, training is usually necessary to enable all coders to rely upon the same aspects of their experience in their decisions. Experimental studies have demonstrated that training prior to coding can significantly increase the level of intercoder agreement (Kaplan and Goldsen, 1949; Woodward and Franzen, 1948). In another study, however, nondirected discussion of categories and rules failed to result in significantly higher agreement (Spiegelman, Terwilliger, and Fearing, 1953b).

Category reliability

One goal of a content-analysis research design is to formulate categories "for which the empirical evidence is clear enough so that competent judges will agree to a sufficiently high degree on which items of a certain population belong in the category and which do not" (Schutz, 1958, p. 512). In coding content data, the judges must first be able to agree with respect to the boundaries of units coded—sometimes called the process of *unitizing*. If the unit is a symbol, a paragraph, or an item, unitizing presents relatively few difficulties because the data provide coders with certain physical guides. For example, the symbol is bounded by spaces, the paragraph is set off by indentation, and so on. Thematic analysis presents the most serious problem because the theme is not a "natural unit" for which physical guides exist. Many sentences contain more than one theme, and the proper boundaries between them are often ambiguous.

In addition to identifying boundaries of the content unit, the judge must decide the category into which the unit is to be placed. Reliability of classification is largely a function of category definition and the types and numbers of discriminations to be made. Pretesting of categories on a sample of the material to be coded will enable the investigator to determine which categories require further clarification. Guetzkow (1950) has derived reliability estimates for both unitizing and categorizing operations which permit the investigator to determine how much of the body of data needs to be cross-checked to ensure any desired level of accuracy.

The investigator faced with low agreement levels at the pretesting stage may attempt to solve his problem in one of several ways. Training coders has already been discussed. One might suspect, however, that a major source of error lies in the categories themselves; that untrained coders are more likely to agree if categories are defined clearly than are well-trained coders working with ambiguous categories. This reasoning was supported by at least one content-analysis experiment; for both theme analysis and symbol coding, significant error variance was found between categories, but not between coders (Stempel, 1955).

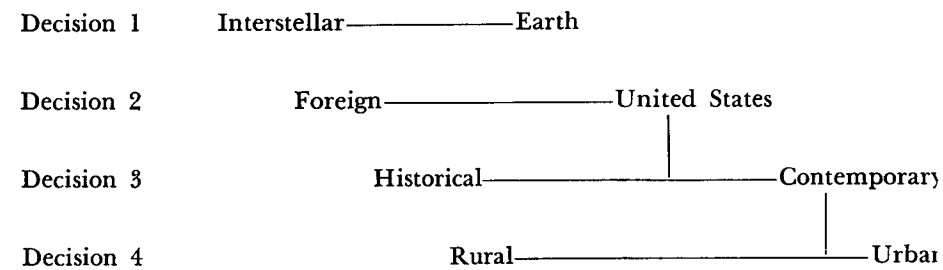
The content-analysis literature contains a number of approaches which may be used to resolve problems of reliability attributable to categories. First, the analyst may define his categories exhaustively, attempting to reduce coding from a judgmental task to a clerical one. An extensive experiment has demonstrated that flexible coding of symbols yields significantly less reliable results than methods in which every member of a category is specified (Geller, Kaplan, and Lasswell, 1942b). However, inasmuch as few categories lend themselves to exhaustive definition, this solution is appropriate only for a limited number of research problems.

Second, fine discriminations between categories often result in a high incidence of disagreement. After pretesting, the investigator may aggregate such categories, but this approach is applicable only if the fine distinctions are not of major theoretical significance.

A third approach to the problem of low reliability is the introduction of additional judges. While this expedient may be necessary for the most difficult

judgmental tasks—for example, scaling the intensity of the themes (North *et al.*, 1963)—it adds considerably to research costs and is a poor substitute for precise coding rules.

A solution more generally applicable to decisions of categorization—but not unitizing or scaling for intensity—consists of reducing each judgment to a dichotomous decision; that is, one in which only a single operation is required of the judge (Schutz, 1958). Consider an example from the study of comic strips. Judges were instructed to code the location of action into six categories: United States, rural, historical, interstellar, urban, or foreign. One source of difficulty in this classification scheme is that the categories are at different levels of generality; for example, the distinction between "United States" and "foreign" actually depends on a prior decision of a more general nature, the choice between "interstellar" and "earth." Arrangement of these categories into a series of dichotomous decisions can be illustrated schematically:



Note that for purposes of this study, not all logically possible categories were used for example, once an item was classified as "foreign," no further discrimination as to "historical-contemporary" or "rural-urban" was required. Schutz (1958 p. 507) found that interjudge agreement was raised from 61 percent with traditional methods of judgment to not less than 90 percent with the dichotomous-decision method. Tables of confidence limits have been developed, from which one can determine, knowing the number of judges and the number of judgments they are required to make, the percentage of agreement necessary for acceptance at the 0.01 and 0.05 levels (Schutz, 1952).

There are several advantages associated with the dichotomous-decision technique. First, it permits coders to focus on a single decision at a time, and to review the criteria for choice at each step. It has been demonstrated that, with traditional methods, increasing the number of categories within the category set decreases reliability (Janis, Fadner, and Janowitz, 1943). Thus the dichotomous decision method should be particularly useful when many categories are necessary. Second, difficulties arise when the process of categorization consists of several judgments, but one decision is logically prior to another because it is relevant to a larger class, as in the example cited above from the study of comic strips. The dichotomous-decision method ensures that the choices given to judges are logical. Third, the method permits the analyst to determine precisely where agreement

between judges is breaking down, information which is useful in redefining categories.

Intercoder agreement may be computed by a variety of methods, of which correlation and percentage of agreement are most frequently used. Simple percentage of agreement is not an adequate measure of reliability because it does not take into account the extent of intercoder agreement which may result from chance (Bennett, Alpert, and Goldstein, 1954). By chance alone agreement should increase as the number of categories decreases. Scott (1955) subsequently developed an index of reliability (π) which corrects not only for the number of categories in the category set, but also for the probable frequency with which each is used. Originally developed for coding data into nominal categories, Scott's formula may also be used for ordinal and interval data. (For an extensive illustration of this method, see Angell, Dunham, and Singer, 1964). Other methods of computing agreement applicable to content analysis are described in Kaplan and Goldsen (1949), Robinson (1957), and Stempel (1955), but Scott's formula, which produces a conservative estimate of reliability, appears to be the most useful.

A method of computing and reporting the overall reliability of any category set with a single figure has been developed (Spiegelman, Terwilliger, and Fearing, 1953b). The technique involves ranking the patterns of agreement among judges on each item. For example, with four judges, five patterns are possible: 4 (complete agreement), 3-1, 2-2, 2-1-1, and 1-1-1-1 (complete disagreement). The reliability of the category set is then reported as the mean rank order of all items. With four judges, scores would range from 1.0 (complete agreement on all items) to 5.0 (complete disagreement on all items).

Although individual and category reliability have been discussed separately, this distinction is a somewhat artificial one. Indeed, the investigator may find that it is not always readily apparent to which factor low reliability should be attributed. The "Random-Systematic-Error Coefficient" is a measure which can be used to determine the nature and source of disagreement in coding. Interpretation of the RSE Coefficient is based on the finding that "errors resulting from a defective code (for example, an ambiguous code) seem generally to be scattered about the range of possible disagreements, while errors originating in the coders tend to fall into systematic patterns" (Funkhouser and Parker, 1966, p. 3). Calculation of the RSE Coefficients can be programmed into the same computer routines that calculate Scott's π for the overall test of reliability.

An *acceptable level of reliability* is one of many issues for which there is no ready definition. The question can only be answered in the context of a given research problem. That high reliability can be achieved for simple forms of content analysis, in which coding is essentially a mechanical task, is amply documented in the literature. Conversely, as categories and units of analysis become more complex, they are likely to become both more useful and less reliable. In formulating and testing a content-analysis research design, the analyst may thus be forced to strike some appropriate balance between reliability and problem significance.

VALIDITY

Validity is often defined as the extent to which an instrument is measuring what it is intended to measure. The validity of any study is inextricably interrelated with its sampling design and reliability. If the analyst studying the proportion of international news in the American daily newspaper samples only issues of the *New York Times*, well-constructed categories or precise measurement will not ensure the validity of his findings. Even if the universe from which the sample was drawn were expanded to make it more representative, low reliability on categorizing international news would render the results suspect. Thus, adequate sampling and reliability are necessary, but not sufficient, conditions for validity.

The meaning of validity may differ from study to study, depending on the investigator's purposes. The American Psychological Association Committee on Psychological Tests has distinguished between *content validity*, *predictive validity*, *concurrent validity*, and *construct validity*. The distinction between them can be illustrated by an example. Documents may be analyzed by means of a measure such as the type-token ratio, the number of different words in message samples of a given length. Results may be used as a direct measure of the author's vocabulary; in this case the analyst is concerned with content validity. The same measure might also be used to predict subjects' success in college; to distinguish schizophrenics from normal persons; or to make inferences about the writer's general intellectual capacity. For these purposes, the investigator would be interested in the predictive validity, concurrent validity, and construct validity, respectively, of his measure (Technical Recommendations . . . , 1954, p. 13).

Content validity, also sometimes referred to as face validity, has been most frequently relied upon by content analysts. If the purpose of the research is purely descriptive one, content validity is normally sufficient. Content validity is usually established through the informed judgment of the investigator—that is, "Are the results plausible?"

Predictive validity is concerned with the ability of an instrument to predict events for which evidence is not currently available. On the basis of his data, the analyst may predict the occurrence of future events, or of events for which data are at present inaccessible. An example of predictive validity has been cited in a study of Knut Hamsun's novels; the inference that Hamsun was a latent fascist was later validated by his collaboration with the Nazis (Lowenthal, 1949). Extensive use of predictive validity in the content-analysis literature exists in the area of propaganda analysis. Access to Nazi documents following World War II permitted those engaged in predicting aspects of Axis behavior to assess the accuracy of their inferences (George, 1959a, pp. 253-284; Berelson and De Grazia, 1947).

Concurrent validity is also established by prediction to an outside criterion; it differs from predictive validity only with respect to the time element. If a measure is able to distinguish sources with known differences, as the type-token ratio has been done with schizophrenic and normal patients, the validity of the measure for that purpose is confirmed. Documents written by designated decision makers during

the crisis leading to World War I were content-analyzed to determine day-to-day changes in international "tensions." In order to test the validity of the content measures, a number of independent indices assumed to be sensitive to international tensions—stock market prices, gold flow, commodity futures, and exchange rates—were charted on a daily basis. A high correlation between financial indices and content data enhanced confidence in the latter measure (Holsti and North, 1966).

The important aspect of both predictive and concurrent validity is the criterion; it may be difficult to ensure that the criterion itself is valid. The authors of a simplified readability formula demonstrated that their test correlated significantly with those derived by the Flesch method (Farr, Jenkins, and Patterson, 1951). To the extent that the latter is a true measure of readability, the validity of the Farr formula was established. But it has been demonstrated that the applicability of the Flesch method is limited to certain types of writing (Taylor, 1953); thus, unless established by other methods, the validity of the Farr formula is similarly limited.

Construct validity is concerned not only with validating the measure, but also the theory underlying the measure. It may be established by several methods; that most frequently discussed in the literature involves interrelating the measure within a "nomological network" of external variables from many different sources (Cronbach and Meehl, 1955, p. 290; see also Janis, 1949, pp. 78–79, and Technical Recommendations . . . , 1954).

Corroborating construct validity in content analysis can be illustrated with the concept of "need achievement." A number of studies have established an impressive network of evidence supporting the validity of both the measure and its underlying theory. That need achievement can successfully differentiate between groups known to differ in regard to relevant characteristics, as predicted by the theory, has been demonstrated in a study of various Indian tribes (McClelland and Friedman, 1952) and nations (McClelland, 1961). Need achievement, as measured by content analysis of diverse literary products, has also been shown to correlate with broad patterns of culture change in settings as different as ancient Greece and the United States. Finally, the concept has been found to correlate significantly with such external measures of entrepreneurial achievement as trade and issuance of patents (McClelland, 1958; deCharms and Moeller, 1962).

These, then, are some of the problems which content analysts share with all scientists. On the whole they have been somewhat lax in this respect; for example, in regard to reliability, Berelson noted in 1952 (p. 172):

Whatever the actual state of reliability in content analysis, the published record is less than satisfactory. Only about 15–20% of the studies report the reliability of the analysis contained in them. In addition, about five reliability experiments are reported in the literature.

The recent literature reveals that content analysts are sharing the widespread concern within the social sciences with the issues of sampling, reliability, and

validity. At the same time, it is clear both that these problems have by no means been resolved, and that progress has been far from even. For example, one has little difficulty finding either experimental studies or imaginative examples of sampling the printed mass media. On the other hand, given the tendency of analysts to draw inferences as to the causes of communication from their data, considerably less than a commensurate share of attention has been paid to problems of validity.

COMPUTERS IN CONTENT ANALYSIS

The most significant recent development in content analysis is the programming of electronic computers to handle a variety of operations involved in textual analysis.

Manual methods of content analysis suffer in varying degrees from a number of limitations. Even elementary forms of the method are expensive and time-consuming. Moreover, most techniques lack flexibility and have a limited ability to deal with complex units. Finally, content analysis usually requires skilled and sensitive coders, the very type of persons who soon become bored and frustrated by the tedious and repetitive nature of the task. These difficulties lend considerable support for Berelson's warning (1952, p. 198):

Unless there is a sensible, or clever, or sound, or revealing, or unusual, or important notion underlying the analysis, it is not worth going through the rigor of the procedure, especially when it is so arduous and so costly of effort.

Many of these problems can be minimized or overcome by use of computers but computers are not currently able to undertake all the repetitive and routine chores associated with content analysis. Nor is use of computers warranted for every type of research. For example, the analyst wishing to determine the space allocated in newspapers to various types of news may find computers of little use; the research could be completed more efficiently using simpler instruments of measurement such as a ruler. In general, content-analysis problems which are most appropriately analyzed by space/time or item measures will profit little from computers, except perhaps in the final stages of research for purely numerical operations (cross tabulations, correlational analyses, and the like). On the other hand, computers can be of significant help in research for which the symbol or the theme is a suitable unit of analysis.

At present, most computer content-analysis programs fall into one of two categories. Those of the first type are essentially word-count programs, the output consisting of the frequency with which each word in the text appears. The second type of computer program is characterized by a dictionary system in which text words are looked up in the dictionary and automatically coded with information representing the investigator's frame of reference and assumptions. The coded text may then be manipulated, categorized, tallied, and retrieved according to the analyst's data requirements.

WORD-COUNT PROGRAMS

Word or symbol counting is one of the most widely used forms of content analysis. Although considerably less complex than some other methods, notably theme analysis, symbol counting can nevertheless present serious problems of reliability. In the RADIR studies, for example, although categories were defined exhaustively, reliability tests on inclusion and exclusion indicated only 66, 68, and 70 percent agreement (Lasswell, Lerner, and Pool, 1952, p. 62). This is one of the operations for which computers are ideally suited. Computers will perform frequency counts at high speeds with perfect reliability, provided there are no errors in the punched IBM cards. Moreover, computers can pick up all symbols, not only those believed *a priori* to be of interest or significance. Thus examination of the output may well reveal the appearance of theoretically important symbols which might not have been considered in the original research design.

There are currently several word-counting programs being used across a broad spectrum of research problems. Content analysis based on word counting has been used extensively in psychotherapy. A program for building specialized dictionaries within the vocabulary usage of individual patients has been developed. The computer prints out, in rank order of frequency, the patient's entire vocabulary, as well as tabulating a type-token ratio (Starkweather and Decker, 1964). A somewhat similar set of programs also being used for content analysis of psychotherapeutic interviews, orders and lists all words with exact frequencies. The incidence of each different word within a segment of text is correlated with every other word across all segments, and the resulting matrix can be factor-analyzed (Harway and Iker, 1964; Iker and Harway, 1965; see also Jaffe, 1964).

One of the standard readability formulas has been adapted for computer analysis. Despite difficulties in programming—for example, in defining rules of identification of monosyllables—the program operates at better than 99 percent reliability. At the same time, desire for greater efficiency led the investigators to develop a new readability formula more suitable to the capabilities of the machine (Danielson and Bryan, 1963). A method of computing various scores from the "Cloze" readability formula has also been developed. Scores are tabulated for relative entropy, type-token ratios, and noun/verb ratios (Carstenson and Stolz, 1964).

A combination word-count and information-retrieval system produces concordances of poetry (Painter, 1960; Sebeok and Zeps, 1961). Such programs have been used to analyze the complete works of Matthew Arnold and Yeats (Parrish, 1959; Parrish and Painter, 1963). The program first produces an index of all words appearing in the text in the order of frequency. On the concordance pages, the line in which each word appears is followed by the page number in the standard edition of the poet's collected works, the abbreviated title of the poem, and the line number within the poem:

YOUTHFUL

DEAD HER PRINCELY YOUTHFUL HUSBAND . . . 14 CHURCH BROU I 49
LAY BEFORE HIS YOUTHFUL WIFE 14 CHURCH BROU I 50

PASSING ALL HER YOUTHFUL HOUR . . . 131 TRISTAM 1 39
BUT THE BRILLIANT YOUTHFUL KNIGHT . . 134 TRISTAM 1 127
CHATTING WITH HER YOUTHFUL KNIGHT . . 134 TRISTAM 1 214
OF HIS TIMID YOUTHFUL BRIDE 137 TRISTAM 1 269

One limitation of word-count analyses is the problem of context, which may lend a given word a considerably different meaning. This problem is partially solved in a program which searches the text for concepts of importance to the investigator and prints them out together with up to 120 words appearing before and after each key concept. The program may be operated with two search options: exact matching of key and content words, or the matching of initial letters in the key and content words (Danielson and Jackson, 1963).

THE "GENERAL INQUIRER" PROGRAMS

Probably the computer content-analysis programs currently in widest use are those which have been developed as part of the "General Inquirer" system, "a set of computer procedures for processing 'natural text' . . . that locates, counts, and tabulates text characteristics" (Stone, 1964). Originally developed at the Laboratory for Social Relations at Harvard University for studying psychological and sociological materials (Stone *et al.*, 1962; Dunphy, Stone, and Smith, 1965), this system now encompasses a family of dictionaries, data-preparation systems, and data-analysis programs being used in nearly all social sciences.

Dictionaries

The core of each General Inquirer system of content analysis is a dictionary in which each entry word is defined with one or more "tags" representing categories in the investigator's theory. The dictionary provides the vital link between the theoretical formulation of the research problem and the mechanics of analysis. The necessity for developing rigorous rules concerning "tagging" of words, by forcing unstated assumptions into the open for critical scrutiny, is an important check on many theoretical aspects of the project—unambiguous definition of categories, precise delineation of the boundaries between concepts, and internal logic of the research design.

Entry words may be listed in the dictionary in root form, without frequently appearing suffix endings such as *-e*, *-s*, *-es*, *-ed*, *-ing*, *-ion*, and *-ly*. If a word in the text with such a suffix is not found in the dictionary, the computer will automatically remove the suffix and look up the word root. A single dictionary form (*attack*) can pick up all forms (*attack*, *attacks*, *attacked*, *attacking*) appearing in the text. Thus dictionaries of moderate size, in the range of 4000 words, have a much larger effective capacity. Users of various General Inquirer dictionaries report analysis of 92 to 98 percent of the text, excluding proper nouns. Words appearing in the text but not in the dictionary are printed out separately on a "leftover" list, and may later be added to the dictionary if deemed important. A dictionary entry may also specify that the computer check whether certain neighboring words are present in the text; the tags assigned then depend on the idiom found.

More than a dozen General Inquirer dictionaries are currently in operation. The Harvard psychosociological dictionary of some 3500 entries uses 83 tag categories, all but three of which consist of at least 20 words. Two sets of tags are incorporated into the dictionary. Each word is assigned a single "first-order" tag which represents the common or manifest meaning of the word. These are discrete, independent variables and can be treated as such statistically. Fifty-five first-order tags are grouped under eleven major headings: persons, groups, physical objects, physical qualifiers, environments, culture, emotion, thought, evaluation, social-emotional actions, and impersonal actions.

Entry words may also be assigned "second-order" tags which represent their connotative meanings. The second-order tags are not independent variables. The meaning of an entry word may be defined by using as many of the 28 second-order tags—which refer to institutional contexts, status connotations, and psychological themes—as appear necessary to give a satisfactory definition. For example, *teacher* is tagged with three meanings: job-role, higher-status, and academic—one first-order tag followed by two second-order tags.

The Harvard dictionary has been used in a number of different applications, including an examination of self-analytic small groups (Dunphy, 1966), a case study of the "Letters from Jenny" (Paige, 1966), and comparative analyses of Presidential nomination acceptance speeches (Smith, with Stone and Glenn, 1966), writings of popular and unpopular students (Goldberg, 1966), projected autobiographies by Radcliffe and Egyptian students (Dahlberg and Stone, 1966), reports written by successful and unsuccessful volunteers for field work in Africa (Ramallo, 1966), and real and simulated suicide notes (Ogilvie, Stone, and Shneidman, 1966).

The Stanford political dictionary incorporates considerable theoretical work on semantic differentiation (Osgood, 1962; Osgood, Suci, and Tannenbaum, 1957). Nearly 4000 words are tagged along three dimensions—positive-negative, strong-weak, and active-passive. These dimensions correspond to the *evaluative*, *potency*, and *activity* dimensions which have been found to be primary in human cognition in a variety of cultures. Each tag is further defined for three levels of intensity. Sample entries in the dictionary include the following:

ABANDON = NEGATIVE 2 WEAK 3 PASSIVE 3
 ABET = POSITIVE 2 ACTIVE 3
 ABSURD = NEGATIVE 2

In addition to the main dictionary, a dictionary of proper names has been written. Names of persons and places serve as entry words which are tagged with as much information as desired for identification. The following are some examples of current entries:

GROMYKO = SOVIET-UNION FOREIGN-MINISTER EXECUTIVE COMMUNIST
 WASHINGTON = UNITED-STATES CAPITAL
 SOVIET-UNION = NATION EUROPE COMMUNIST
 RUSSIA = NATION EUROPE COMMUNIST SOVIET-UNION

The proper-name dictionary ensures uniformity of identification, and serves as a cross-reference proper names, as with Gromyko, Soviet Union, and Russia in the above example. In addition, this dictionary is useful for analyses which depend on discriminating between units and subunits, or for aggregating subunits. The version of the General Inquirer is being used to develop and test hypotheses relating to decision making in crisis (Holsti, Brody, and North, 1965), cohesion within the Soviet bloc (Holsti, 1965b; Hopmann, 1967), and attitudinal components of neutralism (Choucri, 1967).

The New Mexico anthropological dictionary written by Colby (1966c) for cross-cultural comparison of folktales and projective test materials, incorporates 99 tag categories, including a number of value categories developed by Clyce Kluckhohn. Initial studies with the dictionary have included analyses of folktales from five cultures (Kwakiutl, Egyptian, Eskimo, Indian, and Chinese), and TA protocols of Navajos and Zuni (Colby, 1966a; Colby, Collier, and Postal, 1966; Colby and Menchik, 1964).

A number of special-purpose dictionaries have also been prepared. A "alcohol" dictionary of 95 separate tags has been constructed for testing hypotheses relating themes to cultural uses of alcohol. This dictionary of about 3600 entries and 99 tags has been used to analyze a worldwide sample of folktale collections (Kalin, Davis, and McClelland, 1966). A set of dictionaries has been developed to analyze themes, images, and evaluations associated with both product and corporate images. Some 2500 entries are classified into 70 tag categories, such as product properties, institutional references, and product areas (Stone *et al.*, 1966). The "need achievement" dictionary (D. M. Ogilvie and Louise Woodhead, Harvard) follows the rules developed by McClelland for manual scoring of achievement imagery in projective test materials. The dictionary classifies about 1200 entries into 25 tag categories. Thirty-eight tag concepts suggested by the theoretical work of Talcott Parsons are used to classify about 2400 entries in the political dictionary developed for analyzing lobbying behavior (McPherson, 1966). In the WAI dictionary, developed for analyzing multiple open-ended responses to the question "Who am I," 3000 entry words are used to define 30 tag categories (McLaughlin, 1966). A dictionary for studying survey responses, based on the Harvard dictionary but incorporating adjustments for language used by middle and lower-class subjects has been written by B. Frisbie (University of Chicago). Eight value categories developed in Lasswell's and Kaplan's *Power and Society* (1950) define the tag categories in a political dictionary written at Yale (Peterson and Brewer, n.d.). Finally, dictionaries have been prepared to score "need for affiliation" (J. Williamson, Harvard), modes of reaction to psilocybin (T. Dinkel, Chicago), and Icarian imagery (D. M. Ogilvie and D. C. Dunphy, Harvard).

One of the many advantages of the General Inquirer system is that the dictionaries are basically interchangeable; thus the investigator may run his data on another dictionary tagged for different variables. For example, Presidential nomination acceptance speeches were analyzed on several different dictionaries (Smith, Stone, and Glenn, 1966). In interpreting the results it is, of course, important to be cognizant of the theoretical assumptions underlying the dictionary.

including the premise that it will be used with populations similar to those for which it was constructed.

Data Preparation

The text to be analyzed is punched on IBM cards with as little or as much pre-editing as required by the analyst's problem. Most investigations use text directly transcribed onto IBM cards without any coding. Minimal coding normally involves the separation of complex sentences into one or more themes or "thought sequences," and identification of indefinite terms such as pronouns.

To answer some research questions on the basis of communications content, it may not be enough to know that *X*, *Y*, or both *X* and *Y* occur in a sentence; it may be more important to know the perceived relationship between *X* and *Y*. Consider the following three sentences:

This Soviet decision is a deliberately provocative and unjustified threat to American security.

This American decision is a deliberately provocative and unjustified threat to Soviet security.

This deliberately provocative unjustified decision is a threat to Soviet and American security.

Each theme contains exactly the same words but their meaning is clearly different. Neither word frequency nor contingency analysis can distinguish among these three sentences. To do so requires some form of theme analysis which, in turn, necessitates some prior coding of data. More elaborate coding systems include the identification of syntactical position of key words in the text, and the addition of certain other codes for the theme as a whole (Holsti, 1964; Stone *et al.*, 1962). For example, a sentence might be coded as follows:

CV A free/3 American/3 people/3 must/4 reject/4 all
fatalistic/7 philosophies/7 of history/7.

These codes identify the subject-verb-object (3-4-7) relationship, links between modifiers and referents, time (C = current), and mode of expression (V = imperative). While such operations add to the time and effort required for data preparation, they also permit use of more elaborate analysis programs which may yield data important to the investigator's research design. As automated language-processing routines and syntax-identification programs become available, it will be possible to forgo most if not all precoding. A realistic appraisal of present achievements and future prospects along these lines has been prepared by Simmons (1966).

Data-analysis programs

The General Inquirer system includes a broad spectrum of programs for analyzing text. A *text and tag list* program prints out the text and tags assigned by sentence

in the form of a bilingual book. A *tag tally* program counts words in the text which have been tagged in the dictionary. In addition to raw scores, an index based on the ratio of occurrence of tag words to total words in the document is computed. For rapid visual interpretation of results, tag tallies can be printed in graph form. A separate list of all text words not found in the dictionary is also printed out.

A *question and search* program retrieves, tallies, and prints out all sentences meeting any desired specifications. The analyst may wish to search the text for all sentences containing a certain *text word* or *cluster of words*—for example, all sentences in which "Soviet Union" occurs as the subject and which also contain the words "nuclear" and "weapons." Questions may also retrieve themes in terms of *tags*, with or without specification of intensity or syntax position. The *theme codes* may also be used for retrieval. Any of the question specifications may be used singly or in any desired combination, and up to 100 retrieval questions can be processed at once.

Parsons, Shils, and Olds (1952, p. 57) have pointed out that social objects may be significant as *complexes of qualities* or *complexes of performance*. In the former case the object is considered in terms of its attributes, of *what he is*, whereas in the latter case the object is viewed in terms of *what he does*. A *direct table* program provides answers to both types of questions for any specified attitude object—nations, institutions, groups, persons, concepts, programs, ideologies, and the like—which appear within the data (Armour, 1964; Holsti, 1966). The program operates only on text with syntax codes. Row headings are defined by the dictionary variables and column headings represent attitude objects specified by the investigator.

If the analysts wished to determine Chinese attitudes, as expressed in any given document or set of documents, toward China, the Test-Ban Treaty, and the original signators of that treaty, the information would be printed out, using the Stanford dictionary, in the following format:

0331 Kuo Mo Jo of China 7-26-63 PEKING DOMESTIC SERVICE BROADCAST NUMBER OF WORDS IN TEXT - 3834					
	U.S.	SOV. UN.	GT. BRIT.	CHINA	TEST BAN
POSITIVE AFFECT	10.00	18.00	4.00	54.20	26.00
NEGATIVE AFFECT	260.40	26.00	28.00	33.00	57.00
STRONG	163.40	55.00	12.00	117.00	77.00
WEAK	2.00	10.00	6.00	5.40	26.00
ACTIVE	129.50	26.00	6.00	71.40	57.00
PASSIVE	8.00	6.00	7.00	21.00	31.00

Information at the top of the table identifies the document number (0331), author (Kuo Mo Jo of China), date (July 26, 1963), and source. Cell scores in the table indicate a summary score (frequency × intensity) of expressed attitudes toward each nation and toward the treaty for tag categories in the dictionary. Immedi-

ately following is a table which indicates the precise frequency count for each intensity level (for example, strong 1, strong 2, strong 3, etc.). Separate pairs of tables are produced to score assertions about qualities and performance.

Scores for up to a dozen attitude objects (that is, the column headings) may be analyzed in each pass through the data. Column headings can be changed simply by inserting different control cards; hence, the data may be rerun as often as is necessary to satisfy the requirements of the research problem. A special feature of this program is that the intensity level of words in the dictionary may be adjusted according to the mode of expression; for example, the score of the word "aid" in the sentence "The United States may aid India" can be reduced by a constant to reflect the probabilistic nature of the assertion. Tables are also punched out in card form for direct statistical analysis by computer. This program may be used with any dictionary punched in a standard format.

Users of content analysis often require psychological indexes that will discriminate between two sources. The General Inquirer program has been combined with the Hunt-Hovland "Concept Learner" to produce a program for identifying discriminant functions. The strategy on which the program is based is to search the text for a single concept (or combination of concepts) which discriminates sentences in document *A* from those in document *B*. Failing this, the program will continue the process of subdivision until a subgroup is found where a test does apply, or one of the document sources runs out of sentences (Hunt, Kreuter, and Stone, 1965; Stone and Hunt, 1963).

Another variation of the General Inquirer approach to content analysis involves efforts to duplicate manual scoring methods by constructing rules that enable *the computer* to analyze its own tag applications and to make decisions about the nature of a document on the basis of tag profiles. The following document was scored by the Woodhead-Ogilvie Need-Achievement scoring system:

SENTENCE 1	(TAGS APPLIED)
THE STUDENT IS <u>DREAMING</u> <u>ABOUT BECOMING A GREAT IN-</u> <u>VENTOR.</u>	NEED <u>TO-BE</u> ADJECTIVE-POSITIVE <u>ROLE-POSITIVE</u> SENTENCE SUMMARY = ACHIEVEMENT IMAGERY
SENTENCE 2	
AFTER <u>YEARS OF LABOR</u> THE CRUCIAL MOMENT ARRIVES.	TIME VERB-POSITIVE SENTENCE SUMMARY = UNRELATED IMAGERY
SENTENCE 3	
HE <u>HOPES EVERYTHING</u> WILL <u>PAN OUT PROPERLY.</u>	NEED <u>VERB-POSITIVE ADVERB-</u> <u>POSITIVE</u> SENTENCE SUMMARY = ACHIEVEMENT IMAGERY

SENTENCE 4

BUT THE INVENTION WILL BE VALUES-POSITIVE FAILURE SEN
A FAILURE. TENCE SUMMARY = UNRELATEI
IMAGERY

SENTENCE 5

HE WILL DIE DISCOURAGED. AFFECT-NEGATIVE SENTENCI
SUMMARY = ACHIEVEMENT
IMAGERY

*****SUMMARY***** THIS DOCUMENT CONTAINS ACHIEVE
MENT IMAGERY

Tags correspond to the underlined sentence words. Underlined tag words indicate that a prespecified achievement-related sequence has been matched. For a document to be scored "Achievement Imagery" it must contain at least one sentence that meets the criteria of one or more of the prespecified rules within the program. For example, Sentence 1 contains the sequence: NEED TO-BE ROLE-POSITIVE. This matches an Achievement Imagery rule. A more complex rule is matched in Sentence 5: this rule states that when AFFECT is evident, check the preceding sentence for mention of SUCCESS or FAILURE. If either of these tags is located, the sentence containing reference to AFFECT is scored Achievement Imagery. The computer scoring method correlates well with manual coding.

OTHER COMPUTER CONTENT-ANALYSIS PROGRAMS

One by-product of recent interest in machine translation of text has been computer programs for syntactical, as distinct from semantic, discriminations. These may be used for a variety of content-analysis problems concerning the question "how it is said"; for example, analysis of literary style (Sedelow, Sedelow, and Ruggles, 1964). One difficulty of these programs is that there may be more than one possible translation from a given sentence. This defect for the purpose of translation can be turned into an advantage for the content analyst. Two studies have shown how the ability of the computer, unencumbered by any preconceptions or "cognitive set," may be used to produce all possible interpretations of legal literature. Allen (1963) and Langevin and Owens (1963) both analyzed sections of the Nuclear Test Ban Treaty to determine the degree of structural ambiguity in its text. While not all solutions were meaningful, several sentences yielded more than one reasonable, and substantively significant, interpretation. Such programs appear to have wide potential application, not only for legal analysis but also in many other aspects of communication research.

Content analysis by computer is in its infancy and it appears certain that the various programs described here will be modified, if not superseded, in the near

future. Five periodical publications may be consulted by the reader wishing to inform himself of the rapid developments which are likely to take place: *Behavioral Science* (quarterly, Mental Health Research Institute, University of Michigan), *Current Research and Development in Scientific Documentation* (semiannual, National Science Foundation), *Computer Newsletter* (occasional, Michigan State University), *Computers and the Humanities* (quarterly, Queens College), and *Computer Studies in the Humanities and Verbal Behavior* (quarterly, Universities of Colorado, Kansas, and North Carolina).

GENERAL IMPLICATIONS OF USING COMPUTERS IN CONTENT ANALYSIS

The most apparent characteristic of computers, the ability to analyze text reliably at almost unbelievable speed, requires no further elaboration. Less obvious, but perhaps of greater importance, are the following points.

First, computers impose rigor and discipline on the formulation of research. The investigator using computers for content analysis is forced to make every step of his research design explicit. For example, the dictionary requires an explicit and unambiguous definition of each variable. Similarly, each step in data analysis by computer must be specified with precision. Every analyst approaches his data with a set of assumptions and a theory, however crude or implicit. At minimum, making premises and rules of inference explicit permits informed communication and evaluation of results. Thus, it is not wholly facetious to suggest that all content-analysis research should be designed as if it were to be done by computer.

Second, when data are punched on IBM cards they are amenable to reanalysis as often and for as many different purposes as desired. Traditional methods of content analysis rarely permit the degree of flexibility necessary to exhaust the potential information in one's data. Even the most meticulously coded data can rarely be used later for answering research questions that were not incorporated in the original design. In conventional content-analysis research the analyst almost of necessity instructs coders to prepare the data to yield answers only for the initial theoretical problems. When content analysis is done manually, if a new hypothesis suggests itself after the data have been coded, the investigator often must choose between recoding the data and dropping the new idea because he is "locked in" by his research design. Data on IBM cards, on the other hand, may be rerun to test hypotheses that had not even been considered at the time of data preparation. The investigator's theory and assumptions are built into the dictionary, not into the data, and the dictionary can be expanded in response to new questions. For example, a set of documents was originally prepared to analyze Sino-Soviet-American interaction during a number of recent crisis periods. Later the same materials were rerun for the purpose of testing a quite different hypothesis about the effects of East-West conflict on relations between Moscow and Peking.

Third, use of computers enables the analyst to undertake very complex data manipulations, such as contingency analyses involving numerous variables, which often cannot be done reliably or economically by hand. When computers are used, the problem of scoring reliability is completely resolved; this does not mean, of course, that the investigator can assume the validity of his results.

Fourth, documents punched on IBM cards can readily be reproduced and exchanged between scholars. At the time of this writing, General Inquirer users have prepared over 50 different studies, comprising over six million words on IBM cards. These have been shared informally and will be formally stored and available at the Survey Research Center Archive, University of Michigan.

Fifth, computers can free the scholar from many of the most laborious chores associated with content-analysis research. The computer, like any tool properly used, can enhance the creativity of the scholar by freeing more of his time for those indispensable ingredients of significant research—the original idea, the creative hunch, the insight which is necessary to make "facts" meaningful.

Despite this optimistic appraisal of the implications of computers, it may be well to conclude on a more cautious note. Just as all research does not lend itself to content analysis, not all content analysis should be done by computer. It is important to remain aware of the dangers in what Kaplan (1964, p. 28) called the "law of the instrument," exemplified by the child who, when given a hammer, suddenly discovers that everything needs pounding. Nor should the limitations of computers be overlooked. Bad data are no better for having been analyzed by a computer. Perhaps the single greatest danger in the use of computers is that the investigator may be lulled into accepting the validity of findings without a critical consideration of the steps preceding and following machine processing. Computers cannot save a sloppy research design, nor will they transform a trivial research problem into an important one. The machine output only reflects the skill and insight—or lack thereof—with which the investigator constructed his dictionary and formulated his research design. Some years ago Bernard Berelson (1954, p. 518) wrote, "Content analysis, as a method, has no magical qualities—you rarely get out of it more than you put in, and sometimes you get less. In the last analysis, there is no substitute for a good idea." Development of computer content-analysis programs detracts nothing from the wisdom of that assertion.

REFERENCES

- Abcarian, G., and S. M. Stanage (1965). Alienation and the radical right. *J. Politics*, 27, 776-796.
- Abu-Lughod, I. (1962). International news in the Arabic press: a comparative content analysis. *Publ. Opin. Quart.*, 26, 600-612.
- Albig, W. (1956). The content of radio programs, 1925-1935. In *Modern public opinion*. New York: McGraw-Hill. Pp. 446-450.

- Albrecht, M. C. (1956). Does literature reflect common values? *Amer. sociol. Rev.*, 21, 272-279.
- *Allen, L. E. (1963). Automation: substitute and supplement in legal practice. *Amer. behav. Scientist*, 7, 39-44.
- *Allport, G. W. (1942). *The use of personal documents in psychological science*. Social Science Research Council. Bulletin No. 49.
- (1946). Letters from Jenny. *J. abnorm. soc. Psychol.*, 41, 315-350, 449-480.
- Allport, G. W., and J. Bruner (1940). Fifty years of change in American psychology. *Psychol. Bull.*, 37, 757-776.
- Almond, G. A. (1954). *The appeals of communism*. Princeton: Princeton Univ. Press.
- Andrews, T. G., and Gertrude Muhlhahn (1943). Analysis of congruent idea patterns as a study in personality. *Char. and Pers.*, 12, 101-110.
- Angell, R. C., Vera S. Dunham, and J. D. Singer (1964). Social values and foreign policy attitudes of Soviet and American elites. *J. Confl. Resol.*, 8, 330-491.
- Armour, Anne (1964). A Balgol program for quantitative format in automated content analysis. Stanford University. (Mimeo)
- Arnheim, R. (1944). The world of the daytime serial. In P. F. Lazarsfeld and F. N. Stanton (Eds.), *Radio research: 1942-1943*. New York: Duell, Sloan, and Pearce. Pp. 34-85.
- Aronson, E. (1958). The need for achievement as measured by graphic expression. In J. W. Atkinson (Ed.), *Motives in fantasy, action and society*. Princeton: Van Nostrand. Pp. 249-265.
- Ash, P. (1948). The periodical press and the Taft-Hartley Act. *Publ. Opin. Quart.*, 12, 266-271.
- Asheim, L. (1950). From book to film. In B. Berelson and M. Janowitz (Eds.), *Reader in public opinion and communication*. Glencoe, Ill.: Free Press. Pp. 299-306.
- *Auld, P., Jr., and E. J. Murray (1955). Content analysis studies of psychotherapy. *Psychol. Bull.* 52, 377-395.
- Backman, C. W. (1956). Sampling mass media content: the use of the cluster design. *Amer. sociol. Rev.*, 21, 729-733.
- Badri, M. B., and W. Dennis (1964). Human-figure drawings in relation to modernization in Sudan. *J. Psychol.*, 58, 421-425.
- Bagdikian, B. H. (1959). The newsmagazines. *The New Republic*, 16, Feb. 2, p. 23.
- Baldwin, A. L. (1942). Personality structure analysis: a statistical method for investigating the single personality. *J. abnorm. soc. Psychol.*, 37, 163-183.
- Bales, R. F. (1950). *Interaction process analysis*. Cambridge, Mass: Addison-Wesley.
- †Barcus, F. E. (1959). Communications content: analysis of the research, 1900-1958. Unpublished doctor's dissertation, University of Illinois.
- (1961). A content analysis of trends in Sunday comics, 1900-1959. *Journal Quart.*, 38, 171-180.
- Barghoorn, F. C. (1964). *Soviet foreign propaganda*. Princeton: Princeton Univ. Press.
- Batlin, R. (1954). San Francisco newspapers' campaign coverage: 1896, 1952. *Journal Quart.*, 31, 297-303.
- Bauer, R. A. (1964). The obstinate audience: the influence process from the point of view of social communication. *Amer. Psychologist*, 19, 319-328.
- Becker, H. P. (1930). Distribution of space in the American Journal of Sociology, 1895-1930. *Amer. J. Sociol.*, 36, 461-466.
- (1932). Space apportioned forty-eight topics in the American Journal of Sociology, 1895-1930. *Amer. J. Sociol.*, 38, 71-78.
- Bennett, E. M., R. Alpert, and A. C. Goldstein (1954). Communications through lim response questioning. *Publ. Opin. Quart.*, 18, 303-308.
- Berelson, B. (1942). The effects of print upon public opinion. In D. Waples (Ed.), *Print, radio, and film in a democracy*. Chicago: Univ. of Chicago Press. Pp. 41-65.
- †——— (1952). *Content analysis in communication research*. Glencoe, Ill.: Free Press.
- †——— (1954). Content analysis. In G. Lindzey (Ed.), *Handbook of social psychology*. Cambridge, Mass.: Addison-Wesley. Pp. 488-518.
- Berelson, B., and S. De Grazia (1947). Detecting collaboration in propaganda. *Publ. Opin. Quart.*, 9, 244-253.
- Berelson, B., and Patricia J. Salter (1946). Majority and minority Americans: an analysis of magazine fiction. *Publ. Opin. Quart.*, 10, 168-190.
- Berkman, D. (1963). Advertising in *Ebony* and *Life*: Negro aspirations vs. reality. *Journalism Quart.*, 40, 53-64.
- Bernstein, M. H. (1953). Political ideas of selected American business journals. *Publ. Opin. Quart.*, 17, 258-267.
- Blumberg, N. B. (1954). *One party press? Coverage of the 1952 presidential campaign by 35 daily newspapers*. Lincoln: Univ. of Nebraska Press.
- Boder, D. P. (1940). The adjective-verb quotient: a contribution to the psychology of language. *Psychol. Record*, 3, 310-343.
- Breed, W. (1958). Comparative newspaper handling of the Emmett Till case. *Journalism Quart.*, 35, 291-298.
- Brinegar, C. (1963). Mark Twain and the Quintus Curtius Snodgrass letters: a statistical test of authorship. *J. Amer. Statist. Assoc.*, 58, 85-96.
- Brinton, J., and W. A. Danielson (1958). A factor analysis of language elements affecting readability. *Journalism Quart.*, 35, 420-426.
- Brody, R. A. (1963). Some systemic effects of the spread of nuclear weapons technology: a study through simulation of a multi-nuclear future. *J. Confl. Resol.*, 7, 663-753.
- Brooks, M. R. (1959). *The Negro press re-examined*. Houston: Christopher.
- Broom, L., and Shirley Reece (1955). Political and racial interest: a study in content analysis. *Publ. Opin. Quart.*, 19, 5-19.

* Includes useful special bibliography.

† Includes useful general bibliography.

- Brown, W. B. (1960). *The people's choice*. Baton Rouge: Louisiana State Univ. Press.
- Bruner, J. S. (1941). The dimensions of propaganda: German shortwave broadcasts to America. *J. abnorm. soc. Psychol.*, *41*, 311-337.
- Budd, R. W. (1964). U.S. news in the press down under. *Publ. Opin. Quart.*, *28*, 39-56.
- Budlong, D. H. (1952). Analysis of radio programs by four commentators. *Journalism Quart.*, *29*, 458-459.
- Bush, C. R. (1961). A system of categories for general news content. *Journalism Quart.*, *38*, 312-322.
- Cahnman, W. J. (1948). A note on marriage announcements in the New York Times. *Amer. sociol. Rev.*, *13*, 96-97.
- Carstenson, F. W., and W. S. Stolz (1964). Cloze procedure analysis programs. *Computer Newsletter*, *3*, 5-6.
- Carter, R. E., Jr. (1957). Segregation and the news: a regional content study. *Journalism Quart.*, *34*, 3-18.
- †Cartwright, D. P. (1953). Analysis of qualitative material. In L. Festinger and D. Katz (Eds.), *Research methods in the behavioral sciences*. New York: Holt, Rinehart, and Winston. Pp. 421-470.
- Choucri, Nazli (1967). Nonalignment in international politics: an analysis of attitudes and behavior. Doctoral dissertation, Stanford University.
- Christenson, R. M. (1964). Report on the Reader's Digest. *Columbia Journalism Rev.*, *3*, 30-36.
- Coats, W. J., and S. W. Mulkey (1950). A study in newspaper sampling. *Publ. Opin. Quart.*, *14*, 533-546.
- Colby, B. N. (1966a). The analysis of culture content and the patterning of narrative concern in texts. *Amer. Anthropologist*, *68*, 374-388.
- (1966b). Cultural patterns in narrative. *Science*, *151*, 793-798.
- (1966c). Development and application of an anthropological dictionary. In P. J. Stone, D. C. Dunphy, M. S. Smith, and D. M. Ogilvie, *The General Inquirer: a computer approach to content analysis in the behavioral sciences*. Cambridge: M.I.T. Press.
- Colby, B. N., G. A. Collier, and Susan K. Postal (1963). Comparison of themes in folktales by the General Inquirer system. *J. Amer. Folklore*, *76*, 318-323.
- Colby, B. N., and M. D. Menchik (1964). A study of Thematic Apperception Tests with the General Inquirer system. *El Palacio*, *71*, 29-36.
- Coleman, E. B. (1962). Improving comprehensibility by shortening sentences. *J. appl. Psychol.*, *44*, 131-134.
- Content analysis—a new evidentiary technique (1948). *Univ. Chicago Law Rev.*, *15*, 910-925.
- Cony, E. R. (1953). Conflict-cooperation of five American dailies. *Journalism Quart.*, *30*, 15-22.
- Cornwell, E. E., Jr. (1959). Presidential news: the expanding public image. *Journalism Quart.*, *36*, 275-283.
- *Cronbach, L., and P. Meehl (1955). Construct validity of psychological tests. *Psych. Bull.*, *52*, 281-302.
- Dahlberg, Frances M., and P. J. Stone (1966). Cross-cultural contrasts in interspersed structuring. In P. J. Stone, D. C. Dunphy, M. S. Smith, and D. M. Ogilvie, *The General Inquirer: a computer approach to content analysis in the behavioral sciences*. Cambridge: M.I.T. Press. Pp. 589-602.
- Dale, E. (1935). *The content of motion pictures*. New York: Macmillan.
- Dale, E., and Jeanne S. Chall (1948). A formula for predicting readability. *Educ. Bull.*, *27*, 11-20, 37-54.
- Danielson, W. A., and J. B. Adams (1961). Completeness of press coverage of the 1960 campaign. *Journalism Quart.*, *38*, 441-452.
- Danielson, W. A., and S. D. Bryan (1963). Computer automation of two readability formulas. *Journalism Quart.*, *40*, 201-206.
- Danielson, W. A., and H. Jackson (1963). A computer program for scanning tapes for key concepts and their immediate contexts. *Computer Newsletter*, *1*, 1-2.
- Davis, F. J., and L. W. Turner (1951). Sampling efficiency in quantitative newspaper content analysis. *Publ. Opin. Quart.*, *15*, 762-763.
- deCharms, R., and G. H. Moeller (1962). Values expressed in American children's drawings: 1900-1950. *J. abnorm. soc. Psychol.*, *64*, 136-142.
- De Fleur, M. L. (1962). Mass communication and the study of rumor. *Sociol. Inquiry*, *32*, 51-70.
- (1964). Occupational roles as portrayed on television. *Publ. Opin. Quart.*, *28*, 57-66.
- Deutsch, F. (1947). Artistic expression and neurotic illness. *Amer. Imago*, *4*, 64-102.
- Dibble, V. K. (1963). Four types of inference from documents to events. *History Theory*, *3*, 203-221.
- Dollard, J., and O. H. Mowrer (1947). A method of measuring tension in written documents. *J. abnorm. soc. Psychol.*, *42*, 3-32.
- Dornbusch, S. M., and Lauren C. Hickman (1959). Other-directedness in consumer-goods advertising: a test of Riesman's historical theory. *Soc. Forces*, *38*, 99-102.
- Dovring, Karin (1954). Quantitative semantics in 18th century Sweden. *Publ. Opin. Quart.*, *18*, 389-394.
- (1959). *Road of propaganda: the semantics of biased communication*. New York: Philosophical Library.
- Dunphy, D. C. (1964). Content analysis—development and critical issues. Cambridge: Harvard University. (Mimeo)
- (1966). Social change in self-analytic groups. In P. J. Stone, D. C. Dunphy, M. S. Smith, and D. M. Ogilvie, *The General Inquirer: a computer approach to content analysis in the behavioral sciences*. Cambridge: M.I.T. Press. Pp. 287-340.
- Dunphy, D. C., P. J. Stone, and M. S. Smith (1965). The general inquirer: further developments in a computer system for content analysis of verbal data in the social sciences. *Behav. Sci.*, *10*, 468-480.

- Eckhardt, W. (1965). War propaganda, welfare values and political ideologies. *J. Confl. Resol.*, 9, 345-358.
- *Ekman, P. (1965). Communication through nonverbal behavior: a source of information about an interpersonal relationship. In S. S. Tomkins and C. E. Izard (Eds.), *Affect, cognition and personality*. New York: Springer.
- Elizur, A. (1949). A content analysis of the Rorschach with regard to anxiety and hostility. *Rorschach Res. Exch.*, 13, 247-284.
- Ellegård, A. (1962). *A statistical method for determining authorship*. Göteborg, Sweden: Acta Universitatis Gothoburgensis.
- Ellison, J. W. (1965). Computers and the Testaments. In *Computers for the humanities*. New Haven: Yale Univ. Press. Pp. 64-74.
- Eto, S., and T. Okabe (1965). Content analysis of statements in regard to Japan made by the People's Republic of China. *Developing Economies*, 3, 49-72.
- Exline, R. V., and Barbara H. Long (1965). An application of psychological scaling methods to content analysis: the use of empirically derived criterion weights to improve intercoder reliability. *J. appl. Psychol.*, 49, 142-149.
- Fairbanks, H. (1944). Studies in language behavior: II. The quantitative differentiation of samples of spoken language. *Psychol. Monogr.*, 56, 13-38.
- Farr, J. N., J. J. Jenkins, and D. G. Patterson (1951). Simplification of the Flesch reading ease formula. *J. appl. Psychol.*, 35, 333-337.
- Fillenbaum, S., and L. V. Jones (1962). An application of 'Cloze' technique to the study of aphasic speech. *J. abnorm. soc. Psychol.*, 65, 183-189.
- Flesch, R. (1943). *Marks of readable style*. New York: Teachers College, Columbia Univ.
- (1948). A new readability yardstick. *J. appl. Psychol.*, 32, 221-233.
- Foster, H. S. (1935). How America became belligerent: a quantitative study of war news. *Amer. J. Sociol.*, 40, 464-475.
- (1937). Charting America's news of the world war. *Foreign Affairs*, 15, 311-319.
- Foster, H. S., and C. J. Friedrich (1937). Letters to the editor as a means of measuring the effectiveness of propaganda. *Amer. polit. Sci. Rev.*, 31, 71-79.
- Funkhouser, G. R., and E. B. Parker (1966). A method of analyzing coding reliability. In E. B. Parker et al., *Patterns of adult information seeking*. Stanford University: Institute for Communication Research.
- Galtung, J., and Mari Holmboe Ruge (1965). The structure of foreign news: the presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *J. Peace Res.*, 1, 64-91.
- Gardiner, L. W. (1962). A content analysis of Japanese and American television. *J. Broadcasting*, 6, 45-52.
- Garver, R. A. (1961). Polite propaganda: "USSR" and "American Illustrated." *Journalism Quart.*, 38, 480-484.
- Geller, A., D. Kaplan, and H. D. Lasswell (1942a). The differential use of flexible and rigid procedures of content analysis. Washington, D.C.: Library of Congress, Experimental Division for Study of War-Time Communications, Document No. 12.
- (1942b). An experimental comparison of four ways of coding editorial content. *Journalism Quart.*, 19, 362-370.
- George, A. L. (1959a). *Propaganda analysis*. White Plains, N. Y.: Row, Peterson.
- (1959b). Quantitative and qualitative approaches to content analysis. In I. de Pool (Ed.), *Trends in content analysis*. Urbana: Univ. of Illinois Press. Pp. 7-32.
- Gerbner, G. (1964). Ideological perspectives and political tendencies in news reporting. *Journalism Quart.*, 41, 495-508.
- Gerbner, G., and P. H. Tannenbaum (1960). Regulation of mental illness content in motion pictures and television. *Gazette*, 6, 365-385.
- Gillie, P. A. (1957). A simplified formula for measuring abstraction in writing. *J. a. Psychol.*, 41, 214-217.
- Ginglinger, Genevieve (1955). Basic values in Reader's Digest, Selection and Constitution. *Journalism Quart.*, 32, 56-61.
- Gleser, Goldine C., L. A. Gottschalk, and Kayla J. Springer (1961). An anxiety scale applicable to verbal samples. *Arch. gen. Psychiat.*, 5, 593-605.
- Goldberg, Janice B. (1966). Computer analysis of sentence completions. *J. proj. T. Pers. Assess.*, 30, 37-45.
- Gordon, D. A. (1952). Methodology in the study of art evaluation. *J. Aesthet.*, 10, 338-345.
- *Gottschalk, L. A., Ed. (1961). *Comparative psycholinguistic analysis of two psychotherapeutic interviews*. New York: International Univ. Press.
- Gottschalk, L. A., and Goldine C. Gleser (1960). An analysis of the verbal content of suicide notes. *Brit. J. Med. Psychol.*, 33, 195-204.
- Gottschalk, L. A., Goldine C. Gleser, E. B. Magliocco, and T. D'Zmura (1961). Further studies on the speech patterns of schizophrenic patients: measuring inter-individual differences in relative degree of personal disorganization and social alienation. *J. ment. Dis.*, 132, 101-113.
- Gottschalk, L. A., Goldine C. Gleser, and Kayla J. Springer (1963). Three hostility scales applicable to verbal samples. *Arch. gen. Psychiat.*, 9, 254-279.
- Gray, W. S., and Bernice E. Leary (1935). *What makes a book readable*. Chicago: Univ. of Chicago Press.
- Guetzkow, H. (1950). Unitizing and categorizing problems in coding qualitative material. *J. clin. Psychol.*, 6, 47-58.
- Hafner, A. J., and A. M. Kaplan (1960). Hostility content analysis of the Rorschach TAT. *J. proj. Tech.*, 24, 137-143.
- Hage, G. S. (1959). Anti-intellectualism in press comment, 1828 and 1952. *Journalism Quart.*, 36, 439-446.
- Haigh, G. (1949). Defensive behavior in client-centered therapy. *J. consult. Psychol.*, 181-189.
- Hall, C. S., and R. L. Van de Castle (1966). *The content analysis of dreams*. New York: Appleton-Century-Crofts.
- Hamilton, T. (1942). Social optimism and pessimism in American Protestantism. *Opin. Quart.*, 6, 280-283.

- Hart, J. A. (1961). The flow of international news into Ohio. *Journalism Quart.*, 38, 541-548.
- (1965). Election campaign coverage in English and U.S. daily newspapers. *Journalism Quart.*, 42, 213-218.
- (1966). Foreign news in U.S. and English daily newspapers: a comparison. *Journalism Quart.*, 43, 443-448.
- Harvey, J. (1953). The content characteristics of best-selling novels. *Publ. Opin. Quart.*, 17, 91-114.
- Harway, N. I., and H. P. Iker (1964). Computer analysis of content in psychotherapy. *Psychol. Reports*, 14, 720-722.
- Haskins, J. B. (1960). Validation of the abstraction index as a tool for counter-effects analysis and content analysis. *J. appl. Psychol.*, 44, 102-106.
- Hatch, D. L., and Mary Hatch (1947). Criteria of social status as derived from marriage announcements in the New York Times. *Amer. sociol. Rev.*, 12, 396-403.
- Hayworth, D. (1930). An analysis of speeches in the presidential campaigns from 1884-1920. *Quart. J. Speech*, 16, 35-42.
- Head, S. W. (1954). Content analysis of television drama programs. *Quart. of Film, Radio and Television*, 9, 175-194.
- Herma, H., H. Kris, and J. Shor (1943). Freud's theory of the dream in American textbooks. *J. abnorm. soc. Psychol.*, 38, 319-334.
- Holsti, O. R. (1964). An adaptation of the 'General Inquirer' for the systematic analysis of political documents. *Behav. Sci.*, 9, 382-388.
- (1965a). East-West conflict and Sino-Soviet relations. *J. appl. behav. Sci.*, 1, 115-130.
- (1965b). The 1914 case. *Amer. polit. Sci. Rev.*, 59, 365-378.
- (1966). Content analysis research in the social sciences. Paper read at IBM-Texas A and M Conference on Computers in Humanistic Research.
- (1967). Cognitive dynamics and images of the enemy. In D. J. Finlay, O. R. Holsti, and R. R. Fagen, *Enemies in politics*. Chicago: Rand McNally. Pp. 25-96.
- Holsti, O. R., R. A. Brody, and R. C. North (1965). Measuring affect and action in international reaction models: empirical materials from the 1962 Cuban crisis. *Peace Res. Soc. Papers*, 2, 170-190.
- Holsti, O. R., and R. C. North (1965). History of human conflict. In E. B. McNeil (Ed.), *Social science and human conflict*. Englewood Cliffs, N.J.: Prentice-Hall. Pp. 155-171.
- (1966). Perceptions of hostility and economic variables. In R. Merritt and S. Rokkan (Eds.), *Comparing nations*. New Haven: Yale Univ. Press. Pp. 169-190.
- Holsti, O. R., R. C. North, and R. A. Brody (1968). Perception and action in the 1914 crisis. In J. D. Singer (Ed.), *Quantitative international politics: insights and evidence*. New York: Free Press. Pp. 123-158.
- Honigfeld, G., A. Platz, and R. D. Gillis (1964). Verbal style and personality authoritarianism. *J. Commun.*, 14, 215-218.

- Hopmann, P. T. (1967). International conflict and cohesion in the communist system. *Int. Stud. Quart.*, 11, 212-236.
- Hovland, C. I., I. L. Janis, and H. H. Kelley (1953). *Communication and persuasion*. New Haven: Yale Univ. Press.
- Hughes, F. (1950). *Prejudice and the press*. New York: Devin-Adair.
- Hunt, E. B., J. Kreuter, and P. J. Stone (1965). *Experiments in induction*. New York: Academic Press.
- Iker, H. P., and N. I. Harway (1965). A computer approach toward the analysis of content. *Behav. Sci.*, 10, 173-182.
- Inkeles, A. (1952). Soviet reactions to the Voice of America. *Publ. Opin. Quart.*, 16, 612-617.
- Inkeles, A., and H. Geiger (1952). Critical letters to the editors of the Soviet press: I. Areas and modes of complaint. *Amer. sociol. Rev.*, 17, 694-703.
- (1953). Critical letters to the editors of the Soviet press: II. Social characteristics and interrelations of critics and the criticized. *Amer. sociol. Rev.*, 18, 12-22.
- Jacob, P. E. (1942). Atrocity propaganda. In H. L. Childs and J. B. Whiton (Eds.), *Propaganda by short wave*. Princeton: Princeton Univ. Press. Pp. 211-259.
- *Jaffe, J. (1964). Computer analysis of verbal behavior in psychiatric interviews. In D. M. Rioch and E. A. Weinstein (Eds.), *Disorders of communication*. Research Publications of the Association for Research on Nervous and Mental Disorders, No. 42. Pp. 389-399.
- *——— (1966). The study of language in psychiatry: psycholinguistics and computational linguistics. In S. Arieti (Ed.), *American handbook of psychiatry*. Vol. 3. New York: Basic Books. Pp. 689-704.
- Janis, I. L. (1943). Meaning and the study of symbolic behavior. *Psychiatry*, 6, 425-439.
- (1949). The problem of validating content analysis. In H. D. Lasswell, N. Leite et al. (Eds.), *The language of politics: studies in quantitative semantics*. New York: George Stewart. Pp. 55-82.
- Janis, I. L., and R. Fadner (1949). The coefficient of imbalance. In H. D. Lasswell, N. Leites et al. (Eds.), *The language of politics: studies in quantitative semantics*. New York: George Stewart. Pp. 155-169.
- Janis, I., R. H. Fadner, and M. Janowitz (1943). The reliability of a content analysis technique. *Publ. Opin. Quart.*, 7, 293-296.
- Janowitz, M. (1944). The technique of propaganda for reaction: Gerald L. Smith's radio speeches. *Publ. Opin. Quart.*, 8, 84-93.
- Johns-Heine, P., and H. H. Gerth (1949). Values in mass periodical fiction, 1921-1940. *Publ. Opin. Quart.*, 13, 105-113.
- Johnson, W. (1944). Studies in language behavior: I. A program of research. *Psychol. Monogr.*, 56, No. 2 (whole No. 255).
- Jones, Dorothy B. (1942). Quantitative analysis of motion picture content. *Publ. Opin. Quart.*, 6, 411-428.

- Jones, R., and R. Carter (1959). Some procedures for estimating 'news hole' in content analysis. *Publ. Opin. Quart.*, 23, 399-403.
- Josselson, H. H. (1953). *The Russian word count*. Detroit: Wayne State Univ. Press.
- Kalin, R., W. N. Davis, and D. C. McClelland (1966). The relationship between use of alcohol and thematic content of folktales in primitive societies. In P. J. Stone, D. C. Dunphy, M. S. Smith, and D. M. Ogilvie, *The General Inquirer: a computer approach to content analysis in the behavioral sciences*. Cambridge: M.I.T. Press. Pp. 569-588.
- Kanzer, M. (1948). Dostoevsky's matricidal impulses. *Psychoanal. Rev.*, 35, 115-125.
- Kaplan, A. (1943). Content analysis and the theory of signs. *Philos. Sci.*, 10, 230-247.
- (1964). *The conduct of inquiry*. San Francisco: Chandler.
- Kaplan, A., and J. M. Goldsen (1949). The reliability of content analysis categories. In H. D. Lasswell, N. Leites et al. (Eds.), *The language of politics: studies in quantitative semantics*. New York: George Stewart. Pp. 83-112.
- Kauffman, P. E., and V. C. Raimy (1949). Two methods of assessing therapeutic progress. *J. abnorm. soc. Psychol.*, 44, 379-385.
- Kayser, J. (1953). *One week's news: comparative study of seventeen major dailies for a seven-day period*. Paris: Paul Dupont, for UNESCO.
- Kearl, B. (1948). A closer look at readability formulas. *Journalism Quart.*, 25, 344-348.
- Kingsbury, Susan M., H. Hart, et al. (1937). *Newspapers and the news*. New York: Putnam.
- Klapper, J. (1960). *The effects of mass communication*. Glencoe, Ill.: Free Press.
- Klapper, J., and C. Glock (1949). Trial by newspaper. *Sci. Amer.*, 180 (February, 1949), 16-21.
- Klein, M. W., and N. Maccoby (1954). Newspaper objectivity in the 1952 campaign. *Journalism Quart.*, 31, 285-296.
- Knepprath, H. E. (1962). The elements of persuasion in the nationally broadcast speeches of Eisenhower and Stevenson during the 1956 presidential campaign. Unpublished doctor's dissertation, University of Wisconsin.
- Kobre, S. (1953). How Florida dailies handled the 1952 presidential campaign. *Journalism Quart.*, 30, 163-169.
- Kracauer, S. (1952). The challenge of qualitative content analysis. *Publ. Opin. Quart.*, 16, 631-642.
- Krippendorff, K. (1966). Content analysis: history and critical issues. Unpublished doctor's dissertation, University of Pennsylvania.
- Kris, E., and N. Leites (1953). Trends in twentieth century propaganda. In B. Berelson and M. Janowitz (Eds.), *Public opinion and communication*. Glencoe, Ill.: Free Press. Pp. 278-288.
- Kuhn, A. (1963). *The study of society: a unified approach*. Homewood, Ill.: Dorsey.
- Laffal, J. (1965). *Pathological and normal language*. New York: Atherton.
- Lane, R. E. (1951). Government regulation and the business mind. *Amer. sociol. Rev.*, 16, 163-173.

- Lang, Gladys E., and K. Lang (1955). The inferential structure of political communications: a study in unwitting bias. *Publ. Opin. Quart.*, 19, 168-183.
- Langevin, R. A., and M. Owens (1963). Application of automatic syntactic analysis to the Nuclear Test Ban Treaty. Burlington, Mass.: Technical Operations Research (Mimeo)
- Larson, O. N., L. N. Gray, and J. G. Fortis (1963). Goals and goal-achievement method in television content: models for anomie? *Sociol. Inquiry*, 33, 180-196.
- Lasswell, H. D. (1927). *Propaganda technique in the world war*. New York: Knopf.
- (1935). *World politics and personal insecurity*. New York: McGraw-Hill.
- (1938). A provisional classification of symbol data. *Psychiatry*, 1, 197-204.
- (1941). The world attention survey. *Publ. Opin. Quart.*, 5, 456-462.
- (1942). Communications research and politics. In D. Waples (Ed.), *Print, radio and film in a democracy*. Chicago: Univ. of Chicago Press. Pp. 101-117.
- †——— (1946). Describing the content of communication. In B. L. Smith, H. D. Lasswell and R. D. Casey (Eds.), *Propaganda, communication and public opinion*. Princeton Princeton Univ. Press. Pp. 74-94.
- (1949). Detection: propaganda detection and the courts. In H. D. Lasswell N. Leites et al. (Eds.), *The language of politics: studies in quantitative semantics*. New York: George Stewart. Pp. 173-232.
- Lasswell, H. D., and Dorothy Blumenstock (1939). *World revolutionary propaganda: a Chicago study*. New York: Knopf.
- Lasswell, H. D., and A. Kaplan (1950). *Power and society*. New Haven: Yale Univ. Press.
- Lasswell, H. D., N. Leites et al. (1949). *The language of politics: studies in quantitative semantics*. New York: George Stewart.
- Lasswell, H. D., D. Lerner, and I. de S. Pool (1952). *The comparative study of symbols*. Stanford: Stanford Univ. Press.
- Lazarsfeld, P. F., and A. H. Barton (1951). Qualitative measurement in the social sciences, classification, typologies, and indices. In D. Lerner and H. D. Lasswell (Eds.), *The policy sciences: recent developments in scope and method*. Stanford: Stanford Univ. Press. Pp. 180-188.
- Lazarsfeld, P. F., B. Berelson, and Hazel Gaudet (1944). *The people's choice: how the voter makes up his mind in a presidential campaign*. New York: Duell, Sloan, and Pearce.
- Leary, T. (1957). *Interpersonal diagnosis of personality*. New York: Ronald.
- Lee, A. M. (1952). *How to understand propaganda*. New York: Rinehart.
- Lee, A. M., and Elizabeth B. Lee, Eds. (1939). *The fine art of propaganda*. New York: Harcourt, Brace.
- Leites, N. C. (1949). Interaction: the Third International on its change of policy. In H. D. Lasswell, N. Leites et al. (Eds.), *The language of politics: studies in quantitative semantics*. New York: George Stewart. Pp. 298-333.
- Leites, N. C., and I. de S. Pool (1942). On content analysis. Washington, D.C.: Library of Congress, Experimental Division for Study of War-Time Communications, Document No. 26.

- (1949). Interaction: the response of communist propaganda to frustration. In H. D. Lasswell, N. Leites *et al.* (Eds.), *The language of politics: studies in quantitative semantics*. New York: George Stewart. Pp. 334–381.
- Lerner, D. (1950). "The American Soldier" and the public. In R. Merton and P. Lazarsfeld (Eds.), *Continuities in social research*. Glencoe, Ill.: Free Press. Pp. 212–251.
- Lewin, H. S. (1947). Hitler youth and the Boy Scouts of America: a comparison of aims. *Hum. Relat.*, 1, 206–227.
- Lewis, H. L. (1960). The Cuban revolt story: AP, UPI and 3 papers. *Journalism Quart.*, 37, 573–578, 646.
- Lindner, R. M. (1950). The content analysis of the Rorschach protocol. In L. Abt and L. Bellak (Eds.), *Projective psychology*. New York: Knopf. Pp. 75–90.
- Lorge, I. (1944). Predicting readability. *Teachers College Record*, 45, 404–419.
- Lowenthal, L. (1944). Biographies in popular magazines. In P. F. Lazarsfeld and F. N. Stanton (Eds.), *Radio research 1942–1943*. New York: Duell, Sloan, and Pearce. Pp. 507–548.
- (1949). The sociology of literature. In W. Schramm (Ed.), *Communications in modern society*. Urbana: Univ. of Illinois Press. Pp. 82–100.
- Lynch, M. D., and A. Effendi (1964). Editorial treatment of India in the New York Times. *Journalism Quart.*, 41, 430–432.
- *McClelland, D. C. (1958). The use of measures of human motivation in the study of society. In J. W. Atkinson (Ed.), *Motives in fantasy, action and society*. Princeton: Van Nostrand. Pp. 518–552.
- *——— (1961). *The achieving society*. Princeton: Van Nostrand.
- McClelland, D. C., and G. A. Friedman (1952). A cross-cultural study of the relationship between child-rearing practices and achievement motivation appearing in folk tales. In G. E. Swanson, T. M. Newcomb, and E. L. Hartley (Eds.), *Readings in social psychology* (2nd ed.). New York: Henry Holt. Pp. 243–249.
- Maccoby, N., F. O. Sabghir, and B. Cushing (1950). A method for the analysis of the news coverage of industry. *Publ. Opin. Quart.*, 14, 753–758.
- McCurdy, H. G. (1939). Literature and personality: analysis of the novels of D. H. Lawrence. *Char. and Pers.*, 8, 181–203, 311–322.
- (1947). A study of the novels of Charlotte and Emily Brontë as an expression of their personalities. *J. Pers.*, 16, 109–152.
- (1953). *The personality of Shakespeare: a venture in psychological method*. New Haven: Yale Univ. Press.
- McDiarmid, M. (1937). Presidential inaugural addresses: a study of verbal symbols. *Publ. Opin. Quart.*, 1, 79–82.
- McEvoy, J., with the collaboration of R. Schmuck and M. Chesler (1966). Letters from the right: content-analysis of a letter writing campaign. Institute for Social Research, University of Michigan. (Mimeo)
- McGranahan, D. V., and I. Wayne (1948). German and American traits reflected in popular drama. *Hum. Relat.*, 1, 429–455.
- McLaughlin, B. (1966). The WAI dictionary and self-perceived identity in college students. In P. J. Stone, D. C. Dunphy, M. S. Smith, and D. M. Ogilvie, *The General Inquirer: a computer approach to content analysis in the behavioral sciences*. Cambridge: M.I.T. Press.
- McPherson, W. (1964). Lobbying and communication processes. Paper read at American Political Science Association Meeting, Chicago.
- Mahl, G. F. (1959). Exploring emotional states by content analysis. In I. de S. Pool (Ed.), *Trends in content analysis*. Urbana: Univ. of Illinois Press. Pp. 89–130.
- Mann, Mary B. (1944). Studies in language behavior: III. The qualitative differentiation of samples of written language. *Psychol. Monogr.*, 56, No. 2 (whole No. 255).
- Mardsen, G. (1965). Content analysis studies of therapeutic interviews: 1954 to 1964. *Psychol. Bull.*, 63, 298–321.
- Markham, J. W., and G. H. Stempel III (1957). Analysis of techniques in measuring press performance. *Journalism Quart.*, 34, 187–190.
- Massing, P. W. (1963). The image of the Voice of America as drawn in Soviet media. In Matilda W. Riley (Ed.), *Sociological research: a case approach*. New York: Harcourt, Brace, and World. Pp. 308–314.
- Matthews, B. C. (1910). A study of a New York daily. *Independent*, 68, 82–86.
- Merrill, J. C. (1962). The image of the United States in ten Mexican dailies. *Journalism Quart.*, 39, 203–209.
- (1965). How Time stereotyped three U.S. Presidents. *Journalism Quart.*, 42, 563–570.
- Merritt, R. L. (1966). *Symbols of American community, 1735–1775*. New Haven: Yale Univ. Press.
- Mendenhall, T. C. (1887). The characteristic curves of composition. *Science*, 9, 237–246.
- Merton, R. K. (1946). *Mass persuasion: the social psychology of a war bond drive*. New York and London: Harper.
- (1957). Science and economy of 17th century England. In R. K. Merton, *Social theory and social structure*. Glencoe, Ill.: Free Press. Pp. 607–627.
- Miles, Josephine (1951). The continuity of English poetic language. Univ. of California Publications in English. Berkeley: Univ. of California Press. Pp. 517–535.
- Mills, T. M. (1964). *Group transformation*. Englewood Cliffs, N.J.: Prentice-Hall.
- Mintz, A. (1949). The feasibility of the use of samples in content analysis. In H. D. Lasswell, N. Leites *et al.* (Eds.), *The language of politics: studies in quantitative semantics*. New York: George Stewart. Pp. 127–152.
- Morris, C. W. (1946). *Signs, language, and behavior*. New York: Prentice-Hall.
- Morton, A. Q. (1963). A computer challenges the church. *The Observer*, Nov. 3, 1963.
- Mosteller, F., and D. L. Wallace (1964). *Inference and disputed authorship: the Federalist*. Reading, Mass.: Addison-Wesley.
- Mott, F. L. (1942). Trends in newspaper content. *Annals*, 219, 60–65.

- Mowrer, O. H. (1953). Changes in verbal behavior during psychotherapy. In O. H. Mowrer (Ed.), *Psychotherapy theory and research*. New York: Ronald. Pp. 463-545.
- Nixon, R. B., and R. R. Jones (1956). The content of non-competitive newspapers. *Journalism Quart.*, 33, 299-314.
- †North, R. C., O. R. Holsti, M. G. Zaninovich, and Dina A. Zinnes (1963). *Content analysis: a handbook with applications for the study of international crisis*. Evanston, Ill.: Northwestern Univ. Press.
- Nunnally, J. (1957). The communication of mental health information: a comparison of the opinions of experts and the public with mass media presentations. *Behav. Sci.*, 2, 222-230.
- Ogilvie, D. M., P. J. Stone, and E. S. Shneidman (1966). Some characteristics of genuine versus simulated suicide notes. In P. J. Stone, D. C. Dunphy, M. S. Smith, and D. M. Ogilvie, *The General Inquirer: a computer approach to content analysis in the behavioral sciences*. Cambridge: M.I.T. Press. Pp. 527-535.
- Ohlström, B. (1966). Information and propaganda. *J. Peace Res.*, 1, 75-88.
- Ojemann, R. H., et al. (1948). A functional analysis of child development material in current newspapers and magazines. *Child Developmt.*, 19, 76-92.
- Osgood, C. E. (1959). The representational model and relevant research methods. In I. de S. Pool (Ed.), *Trends in content analysis*. Urbana: Univ. of Illinois Press. Pp. 33-88.
- (1962). Studies on the generality of affective meaning systems. *Amer. Psychologist*, 17, 10-28.
- Osgood, C. E., S. Saporita, and J. C. Nunnally (1956). Evaluative assertion analysis. *Litera*, 3, 47-102.
- Osgood, C. E., and T. A. Sebeok, Eds. (1954). Psycholinguistics: a survey of theory and research problems. *J. abnorm. soc. Psychol. (Suppl.)*, 49, No. 4, part 2.
- Osgood, C. E., G. J. Suci, and P. H. Tannenbaum (1957). *The measurement of meaning*. Urbana: Univ. of Illinois Press.
- Osgood, C. E., and Evelyn G. Walker (1959). Motivation and language behavior: content analysis of suicide notes. *J. abnorm. soc. Psychol.*, 59, 58-67.
- Page, H. W. (1953). An assessment of the predictive value of certain language measures in psychotherapeutic counseling. In W. U. Snyder (Ed.), *Group report of a program of research in psychotherapy*. University Park, Pa.: Pennsylvania State College. Pp. 88-93.
- Paige, J. M. (1966). Letters from Jenny: an approach to the clinical analysis of personality structure by computer. In P. J. Stone, D. C. Dunphy, M. S. Smith, and D. M. Ogilvie, *The General Inquirer: a computer approach to content analysis in the behavioral sciences*. Cambridge: M.I.T. Press. Pp. 431-451.
- Painter, J. A. (1960). Computer preparation of a poetry concordance. *Commun. Assoc. for Computing Machinery*, 3, 91-95.
- *Paisley, W. J. (1964). Identifying the unknown communicator in painting, literature and music: the significance of minor encoding habits. *J. Commun.*, 14, 219-237.
- (1966). The effects of authorship, topic, structure, and time of composition on letter redundancy in English texts. *J. verb. Learn. verb. Behav.*, 5, 28-34.
- Parker, E. C., D. W. Barry, and D. W. Smythe, Eds. (1955). *The television-radio audience and religion*. New York: Harper.
- Parrish, S. M., Ed. (1959). *A concordance to the poems of Matthew Arnold*. Ithaca, N.Y.: Cornell Univ. Press.
- Parrish, S. M., and J. A. Painter (1963). *A concordance to the poems of W. B. Yeats*. Ithaca, N.Y.: Cornell Univ. Press.
- Parsons, T., E. A. Shils, and J. Olds (1952). Values, motives, and systems of action. In T. Parsons and E. A. Shils, *Toward a general theory of action*. Cambridge: Harvard Univ. Press. Pp. 47-243.
- Peterson, R. L., and T. L. Brewer (n.d.). *The Lasswell value dictionary*. New Haven: Yale University. (Mimeo)
- Pittenger, R. E., C. F. Hockett, and J. J. Danehy (1960). *The first five minutes: a sample of microscopic interview analysis*. Ithaca, N.Y.: Paul Martineau.
- Pool, I. de S. (1951). *Symbols of internationalism*. Stanford: Stanford Univ. Press.
- (1952a). *The 'prestige papers': a survey of their editorials*. Stanford: Stanford Univ. Press.
- (1952b). *Symbols of democracy*. Stanford: Stanford Univ. Press.
- (1954). Symbols, meaning, and social sciences. In L. Bryson (Ed.), *Symbols and values: an initial study*. New York: Harper.
- †———, Ed. (1959). *Trends in content analysis*. Urbana: Univ. of Illinois Press.
- Prothro, J. W. (1956). Verbal shifts in the American presidency: a content analysis. *Amer. polit. Sci. Rev.*, 50, 726-739.
- Raimy, V. C. (1948). Self reference in counseling interviews. *J. consult. Psychol.*, 12, 153-163.
- Rainey, G. E. (1966). *The American image of peace*. Doctoral dissertation, American University.
- Rainoff, T. J. (1929). Wave-like fluctuations of creative productivity in the development of west-European physics in the eighteenth and nineteenth centuries. *Isis*, 12, 287-307.
- Ramallo, L. I. (1966). The integration of subject and object in the context of action: a study of reports written by successful and unsuccessful volunteers for field work in Africa. In P. J. Stone, D. C. Dunphy, M. S. Smith, and D. M. Ogilvie, *The General Inquirer: a computer approach to content analysis in the behavioral sciences*. Cambridge: M.I.T. Press. Pp. 536-547.
- Riesman, D., N. Glazer, and D. Reuel (1950). *The lonely crowd*. New Haven: Yale Univ. Press.
- Roat, R. (1963). Crisis news in the press abroad. *Quill*, June, 1963, pp. 20-21.
- Robinson, W. S. (1957). The statistical measure of agreement. *Amer. sociol. Rev.*, 22, 17-25.
- Role of content analysis in opinion and communications research, The (1951). *Publ. Opin. Quart.*, 15, 782-786.
- Rosen, B. (1964). Attitude change within the Negro press toward segregation and discrimination. *J. soc. Psychol.*, 62, 77-83.

- Rosenau, J. N. (1968). Private preferences and public responsibility: the relative potency of individual and role variables in the behavior of U.S. senators. In J. D. Singer (Ed.), *Quantitative international politics: insights and evidence*. New York: Free Press. Pp. 17-50.
- Roshal, Jean J. G. (1953). The type-token ratio as a measure of changes in behavior variability during psychotherapy. In W. U. Snyder (Ed.), *Group report of a program of research in psychotherapy*. University Park, Pa.: Pennsylvania State College. Pp. 94-104.
- Rosi, E. J. (1964). How 50 periodicals and the Times interpreted the Test Ban controversy. *Journalism Quart.*, 41, 545-556.
- Royal Commission on the Press, 1947-1949 (1949). Report, Appendix VII. London: His Majesty's Stationery Office.
- Rucker, B. W. (1960). News services' crowd reporting in the 1956 presidential campaign. *Journalism Quart.*, 37, 195-198.
- Runion, H. L. (1936). An objective study of the speech style of Woodrow Wilson. *Speech Monogr.*, 3, 75-94.
- Russell, J. T., and Q. Wright (1933). National attitudes on the Far Eastern controversy. *Amer. polit. Sci. Rev.*, 27, 555-576.
- St. George, M., and L. Dennis (1946). *A trial on trial: the great sedition trial of 1944*. National Civil Rights Committee.
- *Salzinger, K., Stephanie Portnoy, and R. S. Feldman (1964). Verbal behavior of schizophrenic and normal subjects. *Ann. N.Y. Acad. Sci.*, 105, 845-860.
- Scheuch, E. K., and P. J. Stone (1964). The General Inquirer approach to an international retrieval system for survey archives. *Amer. behav. Scientist*, 7, 23-28.
- Schneider, L., and S. M. Dornbusch (1958). *Popular religion: inspirational books in America*. Chicago: Univ. of Chicago Press.
- Schorer, M. (1949). Fiction and the 'matrix of analogy.' *Kenyon Rev.*, 11, 539-560.
- Schramm, W., Ed. (1959). *One day in the world's press*. Stanford: Stanford Univ. Press.
- Schubert, G. (1965). Jackson's judicial philosophy: an exploration in value analysis. *Amer. polit. Sci. Rev.*, 59, 940-963.
- Schutz, W. C. (1952). Reliability, ambiguity and content analysis. *Psychol. Rev.*, 59, 119-129.
- (1958). On categorizing qualitative data in content analysis. *Publ. Opin. Quart.*, 22, 503-515.
- Scott, W. A. (1955). Reliability of content analysis: the case of nominal scale coding. *Publ. Opin. Quart.*, 19, 321-325.
- Sebald, H. (1962). Studying national character through comparative content analysis. *Soc. Forces*, 40, 318-322.
- Sebeok, T. A., and V. J. Zeps (1961). Computer research in psycholinguistics: toward an analysis of poetic language. *Behav. Sci.*, 6, 365-369.
- Sedelow, Sally Y., W. A. Sedelow, Jr., and T. Ruggles (1964). Some parameters for computational stylistics: computer aids to the use of traditional categories in stylistic analysis. In *Proceedings of the IBM literary data processing conference*. Yorktown Heights, N.Y.: IBM. Pp. 211-229.
- Shanas, Ethel (1945). The American Journal of Sociology through fifty years. *Amer. J. Sociol.*, 50, 522-533.
- Sheerer, Elizabeth T. (1949). An analysis of the relationship between acceptance of and respect for self and acceptance of and respect for others in ten counseling cases. *J. consult. Psychol.*, 13, 169-175.
- Shneidman, E. S. (1951). *Thematic test analysis*. New York: Grune and Stratton.
- (1961). A psycho-logical analysis of political thinking: the Kennedy-Nixon 'Great Debates' and the Kennedy-Khrushchev 'Grim Debates.' Cambridge: Harvard University (Mimeo)
- *——— (1963). Plan 11. The logic of politics. In L. Arons and M. A. May (Eds.), *Tele vision and human behavior*. New York: Appleton-Century-Crofts. Pp. 177-199.
- Simmons, R. F. (1966). Automated language processing. In C. A. Cuadra (Ed.), *Annual review of information science and technology*. Vol. 1. New York: Interscience. Pp. 137-169.
- Skinner, B. F. (1939). The alliteration in Shakespeare's sonnets: a study in literary behavior. *Psychol. Record*, 3, 186-192.
- Smith, M. S., with P. J. Stone and E. N. Glenn (1966). A content analysis of twenty presidential nominating speeches. In P. J. Stone, D. C. Dunphy, M. S. Smith, and D. M. Ogilvie, *The General Inquirer: a computer approach to content analysis in the behavioral sciences*. Cambridge: M.I.T. Press. Pp. 359-400.
- Smythe, D. W. (1952). Some observations of communications theory. *Audio-visual Commun. Rev.*, 2, 24-37.
- (1953). *Three years of New York television, 1951-1953*. Urbana: National Association of Educational Broadcasters, Univ. of Illinois.
- Snyder, W. E., Ed. (1953). *Group report of a program of research in psychotherapy*. University Park, Pa.: Pennsylvania State College.
- Sorenson, R. C., and T. C. Sorenson (1955). Proposal for the use of content analysis in literary infringement cases. *Soc. Forces*, 33, 262-267.
- Speed, J. G. (1893). Do newspapers now give the news? *The Forum*, 15, 705-711.
- Spiegelman, M., C. Terwilliger, and F. Fearing (1953a). The content of comics: goals and means to goals of comic strip characters. *J. soc. Psychol.*, 37, 189-203.
- (1953b). The reliability of agreement in content analysis. *J. soc. Psychol.*, 3, 175-187.
- Starkweather, J. A. (1956). Content-free speech as a source of information about the speakers. *J. abnorm. soc. Psychol.*, 52, 394-402.
- Starkweather, J. A., and J. B. Decker (1964). Computer analysis of interview center. *Psychol. Reports*, 15, 875-882.
- Stedman, M. S. (1951). Democracy in American communal and socialist literature. *History of Ideas*, 12, 147-154.
- Stempel, G. H., III (1952). Research in brief: sample size for classifying subject matter dailies. *Journalism Quart.*, 29, 333-334.
- (1955). Increasing reliability in content analysis. *Journalism Quart.*, 32, 449-451.

- (1961). The prestige press covers the 1960 presidential campaign. *Journalism Quart.*, 38, 157-163.
- (1965). The prestige press in two presidential elections. *Journalism Quart.*, 42, 15-21.
- Stephenson, W. (1963). Critique of content analysis. *Psychol. Record*, 13, 155-162.
- Stevens, N. E. (1932). The fad as a factor in botanical publication. *Science*, 75, 286-293.
- Stewart, M. (1943). Importance in content analysis: a validity problem. *Journalism Quart.*, 20, 286-293.
- Stock, Dorothy (1949). The self concept and feelings toward others. *J. consult. Psychol.*, 13, 176-180.
- Stone, P. J. (1964). An introduction to the General Inquirer: a computer system for the study of spoken or written material. Harvard Univ. and Simulmatics Corp. (Mimeo)
- Stone, P. J., R. F. Bales, J. Z. Namenwirth, and D. M. Ogilvie (1962). The General Inquirer: a computer system for content analysis and retrieval based on the sentence as a unit of information. *Behav. Sci.*, 7, 484-494.
- Stone, P. J., D. C. Dunphy, and A. Bernstein (1965). Content analysis applications at simulmatics. *Amer. behav. Scientist*, 8, 23-28.
- †Stone, P. J., D. C. Dunphy, M. S. Smith, and D. M. Ogilvie (1966). *The General Inquirer: a computer approach to content analysis in the behavioral sciences*. Cambridge: M.I.T. Press.
- Stone, P. J., and E. B. Hunt (1963). A computer approach to content analysis. In *Proceedings of the spring joint computer conference, 1963*. Pp. 241-256.
- Stone, P. J., D. M. Ogilvie, and D. C. Dunphy (1963). Distinguishing real from simulated suicide notes using General Inquirer procedures. Paper read at the Joint Annual Meeting of the American College of Neuropsychopharmacology.
- Sussman, Leila (1945). Labor in the radio news: an analysis of content. *Journalism Quart.*, 22, 207-214.
- Tauber, Irene B. (1932). Changes in the content and presentation of reading material in Minnesota weekly newspapers, 1860-1929. *Journalism Quart.*, 9, 281-289.
- Tannenbaum, P. H., and B. S. Greenberg (1961a). The effects of bylines on attitude change. *Journalism Quart.*, 38, 535-537.
- (1961b). 'J.Q.' references: a study of professional change. *Journalism Quart.*, 38, 203-207.
- Tannenbaum, P. H., and M. D. Lynch (1960). Sensationalism: the concept and its measurement. *Journalism Quart.*, 37, 381-392.
- Taylor, W. L. (1953). 'Cloze Procedure'; a new tool for measuring readability. *Journalism Quart.*, 30, 415-433.
- (1956). Recent developments in the use of 'Cloze Procedure.' *Journalism Quart.*, 33, 42-48, 99.
- (1957). Gauging the mental health content of the mass media. *Journalism Quart.*, 34, 191-201.

- Technical recommendations for psychological tests and diagnostic techniques (1954) *Psychol. Bull.*, 51, Suppl., 201-238.
- Toch, H. H., S. E. Deutsch, and D. M. Wilkins (1960). The wrath of the bigot: a analysis of protest mail. *Journalism Quart.*, 37, 173-185, 266.
- Walworth, A. (1938). *School histories at war: a study of the treatment of our wars in the secondary school history books of the United States and in those of its former enemies*. Cambridge: Harvard Univ. Press.
- Wang, C. K. A. (1955). Reactions in Communist China: an analysis of letters to newspaper editors. USAF Personnel Training Research Center, Technical Report No. 3 HHRI Project: "Chinese Documents Project." January 1955.
- Waples, D., B. Berelson, and F. R. Bradshaw (1940). *What reading does to people*. Chicago: Univ. of Chicago Press.
- Wayne, I. (1956). American and Soviet themes and values: a content analysis of pictures in popular magazines. *Publ. Opin. Quart.*, 20, 314-320.
- Webb, E. J., and J. R. Salancik (1965). Notes on the sociology of knowledge. *Journalism Quart.*, 42, 591-595.
- Westley, B. H., C. E. Higbie, T. Burke, D. J. Lippert, L. Maurer, and V. A. Stone (1961). The news magazines and the 1960 conventions. *Journalism Quart.*, 40, 525-531, 647.
- White, R. K. (1947). "Black Boy": a value-analysis. *J. abnorm. soc. Psychol.*, 42, 440-461.
- (1949). Hitler, Roosevelt and the nature of war propaganda. *J. abnorm. soc. Psychol.*, 44, 157-174.
- (1951). *Value-analysis: the nature and use of the method*. Glen Gardner, N.J.: Libertarian Press.
- Whiting, A. S. (1960). *China crosses the Yalu*. New York: Macmillan.
- Wilcox, W. (1962). The press of the radical right: an exploratory analysis. *Journalism Quart.*, 39, 152-160.
- Willey, M. (1926). *The country newspaper*. Chapel Hill: Univ. of North Carolina Press.
- Willoughby, W. F. (1955). Are two competing dailies necessarily better than one? *Journalism Quart.*, 32, 197-204.
- Winick, C. (1963). Trends in the occupation of celebrities—a study of news magazine profiles and television interviews. *J. soc. Psychol.*, 60, 301-310.
- Wolfenstein, Martha, and N. Leites (1950). *Movies: a psychological study*. Glencoe, Ill.: Free Press.
- *Wolfson, Rose (1951). Graphology. In H. H. Anderson and G. M. L. Anderson (Eds.) *An introduction to projective techniques*. New York: Prentice-Hall. Pp. 416-456.
- Woodward, J. L. (1930). *Foreign news in American morning newspapers*. New York: Columbia Univ. Press.
- (1934). Quantitative newspaper analysis as a technique of opinion research. *Social Forces*, 12, 526-537.
- Woodward, J. L., and R. Franzen (1948). A study of coding reliability. *Publ. Opin. Quart.*, 12, 253-257.

- Wright, Q., and C. J. Nelson (1939). American attitudes toward Japan and China. *Publ. Opin. Quart.*, 3, 46-62.
- Wyant, Rowena, and Herta Herzog (1941). Voting via the Senate mailbag. *Publ. Opin. Quart.*, 5, 359-382, 590-624.
- Yobson, S., and H. D. Lasswell (1949). Trend: May Day slogans in Soviet Russia. In H. D. Lasswell, N. Leites *et al.* (Eds.), *The language of politics: studies in quantitative semantics*. New York: George Stewart. Pp. 232-297.
- Yule, G. U. (1944). *The statistical study of literary vocabulary*. London: Cambridge Univ. Press.
- Ziminovich, M. G. (1964). An empirical theory of state response: the Sino-Soviet case. Unpublished doctor's dissertation, Stanford University.
- Zinn, Dina A., R. C. North, and H. E. Koch, Jr. (1961). Capability, threat and the outbreak of war. In J. N. Rosenau (Ed.), *International politics and foreign policy*. New York: Free Press. Pp. 469-482.

Methods and Problems in Cross-Cultural Research

JOHN W. M. WHITING, *Harvard University*

The cross-cultural method utilizes data collected by anthropologists concerning the customs and characteristics of various peoples throughout the world to test hypotheses concerning human behavior. Some of the hypotheses tested have been derived from theories of cultural evolution, others from theories concerned with the integration of culture, and still others, particularly in recent years, from theories of individual, developmental, physiological, and social psychology.

Although the comparison of the ways of different peoples is ancient, E. B. Tylor (1889), who presented at a meeting of the Royal Anthropological Institute of Great Britain a paper entitled "On a Method of Investigating the Development of Institutions; Applied to Laws of Marriage and Descent," was the first to use statistical methods for this purpose. It is interesting that Sir Francis Galton, one of the fathers of modern statistics, presided at this meeting. Although the intent of the study was to support Tylor's particular view of cultural evolution, its importance was that in the paper and the discussion which followed, most of the basic assumptions and problems of cross-cultural research were touched upon.

For the next fifty years the method was almost completely neglected, Nieboer's study (1910) of slavery, and Hobhouse, Wheeler, and Ginsberg's (1915) essay on the correlations between material culture and social institutions being notable exceptions. The method was revived by G. P. Murdock (1937) in a test of the evolutionary priority of matrilineal and patrilineal institutions. Since that time, studies using the cross-cultural method have appeared at an ever increasing rate. For a listing of some of these studies see the attached bibliography.

Recently the cross-cultural method has been used more and more by psychologists. In fact, the names of more psychologists appear as authors of the cross-cultural articles listed in the bibliography than of any other discipline. Further-

Wright, Q., and C. J. Nelson (1939). American attitudes toward Japan and China. *Publ. Opin. Quart.*, 3, 46-62.

Wyant, Rowena, and Herta Herzog (1941). Voting via the Senate mailbag. *Publ. Opin. Quart.*, 5, 359-382, 590-624.

Yakobson, S., and H. D. Lasswell (1949). Trend: May Day slogans in Soviet Russia. In H. D. Lasswell, N. Leites *et al.* (Eds.), *The language of politics: studies in quantitative semantics*. New York: George Stewart. Pp. 232-297.

Yule, G. U. (1944). *The statistical study of literary vocabulary*. London: Cambridge Univ. Press.

Zaninovich, M. G. (1964). An empirical theory of state response: the Sino-Soviet case. Unpublished doctor's dissertation, Stanford University.

Zinnes, Dina A., R. C. North, and H. E. Koch, Jr. (1961). Capability, threat and the outbreak of war. In J. N. Rosenau (Ed.), *International politics and foreign policy*. New York: Free Press. Pp. 469-482.

Methods and Problems in Cross-Cultural Research

JOHN W. M. WHITING, *Harvard University*

The cross-cultural method utilizes data collected by anthropologists concerning the customs and characteristics of various peoples throughout the world to test hypotheses concerning human behavior. Some of the hypotheses tested have been derived from theories of cultural evolution, others from theories concerned with the integration of culture, and still others, particularly in recent years, from theories of individual, developmental, physiological, and social psychology.

Although the comparison of the ways of different peoples is ancient, E. F. Tylor (1889), who presented at a meeting of the Royal Anthropological Institute of Great Britain a paper entitled "On a Method of Investigating the Development of Institutions; Applied to Laws of Marriage and Descent," was the first to use statistical methods for this purpose. It is interesting that Sir Francis Galton, one of the fathers of modern statistics, presided at this meeting. Although the intent of the study was to support Tylor's particular view of cultural evolution, its importance was that in the paper and the discussion which followed, most of the basic assumptions and problems of cross-cultural research were touched upon.

For the next fifty years the method was almost completely neglected, Nieboer's study (1910) of slavery, and Hobhouse, Wheeler, and Ginsberg's (1915) essay on the correlations between material culture and social institutions being notable exceptions. The method was revived by G. P. Murdock (1937) in a test of the evolutionary priority of matrilineal and patrilineal institutions. Since that time studies using the cross-cultural method have appeared at an ever increasing rate. For a listing of some of these studies see the attached bibliography.

Recently the cross-cultural method has been used more and more by psychologists. In fact, the names of more psychologists appear as authors of the cross-cultural articles listed in the bibliography than of any other discipline. Further

- Wright, Q., and C. J. Nelson (1939). American attitudes toward Japan and China. *Publ. Opin. Quart.*, 3, 46-62.
- Wyant, Rowena, and Herta Herzog (1941). Voting via the Senate mailbag. *Publ. Opin. Quart.*, 5, 359-382, 590-624.
- Yakobson, S., and H. D. Lasswell (1949). Trend: May Day slogans in Soviet Russia. In H. D. Lasswell, N. Leites *et al.* (Eds.), *The language of politics: studies in quantitative semantics*. New York: George Stewart. Pp. 232-297.
- Yule, G. U. (1944). *The statistical study of literary vocabulary*. London: Cambridge Univ. Press.
- Zaninovich, M. G. (1964). An empirical theory of state response: the Sino-Soviet case. Unpublished doctor's dissertation, Stanford University.
- Zinnes, Dina A., R. C. North, and H. E. Koch, Jr. (1961). Capability, threat and the outbreak of war. In J. N. Rosenau (Ed.), *International politics and foreign policy*. New York: Free Press. Pp. 469-482.

Methods and Problems in Cross-Cultural Research

JOHN W. M. WHITING, *Harvard University*

The cross-cultural method utilizes data collected by anthropologists concerning the customs and characteristics of various peoples throughout the world to test hypotheses concerning human behavior. Some of the hypotheses tested have been derived from theories of cultural evolution, others from theories concerned with the integration of culture, and still others, particularly in recent years, from theories of individual, developmental, physiological, and social psychology.

Although the comparison of the ways of different peoples is ancient, E. Tylor (1889), who presented at a meeting of the Royal Anthropological Institute of Great Britain a paper entitled "On a Method of Investigating the Development of Institutions; Applied to Laws of Marriage and Descent," was the first to use statistical methods for this purpose. It is interesting that Sir Francis Galton, one of the fathers of modern statistics, presided at this meeting. Although the intent of the study was to support Tylor's particular view of cultural evolution, its importance was that in the paper and the discussion which followed, most of the basic assumptions and problems of cross-cultural research were touched upon.

For the next fifty years the method was almost completely neglected. Nieboer's study (1910) of slavery, and Hobhouse, Wheeler, and Ginsberg's (1915) essay on the correlations between material culture and social institutions being notable exceptions. The method was revived by G. P. Murdock (1937) in a test of the evolutionary priority of matrilineal and patrilineal institutions. Since that time, studies using the cross-cultural method have appeared at an ever increasing rate. For a listing of some of these studies see the attached bibliography.

Recently the cross-cultural method has been used more and more by psychologists. In fact, the names of more psychologists appear as authors of the cross-cultural articles listed in the bibliography than of any other discipline. Further