What is Content Analysis?

Many useful definitions of content analysis have been given. An enduringly popular one by Bernard Berelson is as follows: "Content analysis is a research technique for the objective, systematic, and quantitative description of the manifest content of communications." 1

The key words are objective, systematic, quantitative, and manifest. These are the words which distinguish scientific content analysis from the ordinary, informal analysis all of us do every day in reading newspapers or magazines or listening to the conversations of our friends.

Objective means that the categories used to analyse content must be defined so precisely that different persons can analyze the same content using those definitions and get the same results. It also means that highly evaluative categories and terms (good-bad, fair-unfair, beautiful-ugly) are avoided both because these terms are highly subjective and because their meaning tends to change as fashions come and go.

Systematic means that the selection of content to analyze must be based on a formal, predetermined, unbiased plan; in other words, the analyst cannot choose to examine only those elements in the content which happen to fit his hypothesis and ignore all the others. The word removes content analysis, supposedly, from the argumentative, biased collection of data to prove a point.

Quantitative means that the results of the analysis are usually expressed numerically in some way: in frequency distributions, in contingency tables, in correlation coefficients, in ratios and percentages of various sorts. The preference for quantification is understandable; there is simply much more agreement as to what is correct and incorrect procedure within the precise language of mathematics. It is always advisable, of course, to be sure that the assumptions underlying the use of numbers have been met: a fine teasing of statistics is sometimes used to conceal a crumbling analytical cake.

Manifest means that the semantic analysis involved in content analysis is ordinarily of a fairly direct and simple kind: it deals,
as Harold Lasswell, Daniel Lerner and Ithiel de Sola Pool have put it, with "riding on the lines" and not "between the lines." As mentioned earlier, the researcher may be interested in the forces which shaped the message or the effect the message is likely to have, but he does not code the content or store of these latent forces or effects. He codes the content in fairly obvious terms of what it says. This should not be taken to mean that codes always agree about the manifest content of messages. The author has had many battles (particularly, it seems, with women authors) on the proper interpretation of messages. The session, recorded at the faculty women's ten, "What a sweet hat you have on!" manifestly may mean to a male code that the hat is okay, favorable, novel, good. To a female, however, it manifestly may mean that the hat is cheap, dowdy, last year's, and fifteen years too young for the wearer. The point is that manifest is a relative term. It indicates an area of meaning in which content analysts have traditionally tried to operate, rather than a definite point.

The Literature of Content Analysis

The student who decides to use content analysis as a major method of inquiry frequently does so on the basis of an article he has seen in a current journal, say Journalism Quarterly or Public Opinion Quarterly. Too often, however, he thinks the method is new, and he plunges headlong without dipping further into the history of his technique than articles running back four or five years. In doing this, he may be making a serious error. For content analysis, much as it is known today, has a history running back to the 1920's and being about 1925, and no later, he is the father of journalism ever in this century. Hundreds, perhaps thousands, of content analyses have been made since. And although perhaps and theoretical emphases have changed, the serious student can nevertheless gain a needed perspective and a proper feeling for the place of his study in the history of content analysis by carefully examining earlier work. In addition, he may find that some of the trouble-some problems he is encountering have been encountered before and have been solved before; then, he may avoid having to solve them again. In this area, as in many others, he who ignores history is condemned to repeat it.

The studies most commonly concerned with subject matter classifications of newspaper content. An example of a study along those lines, published in reviewing an analysis of political editorials and associate editorials, is in an excellent bit of work on what is the editorial book is read allowed to look individually to ensure Quarterly Journal of Psychology (all may, for example, American Journal of Sociology, The Social Science Press, and others.)

Major Problems of

The method is not a simple technique that can be applied in a straight line. There is a good deal of preliminary work necessary in adapting a method, as in seeing a convincing method, as in the work of the Berkeley Press, "The simplest, the most methodological, is not necessarily the right one, and when it is not employed, a method has no more than was put in, so there is no substitute. Assuming then, the so-called finding, doing it through o
These ideas published in 1959; Dallas Woodworth's analysis of foreign news in morning newspapers. The next major use was in the analysis of political communication. The works of Lasswell and his students and associates are of importance here. A book-length survey of content analysis by Bernard Berelson appeared in 1954. It contains an excellent bibliography of studies up to 1950.\(^4\) Hugh de Sola Pool is the editor of a book published in 1953, *Trends in Content Analysis*, a collection of papers on new approaches to the method. A general emphasis on the use of content analysis in psychology and psycholinguistics is clearly evident. A limited bibliography updates the Berelson book to some extent.\(^5\) The student would probably be well advised to look for recent articles in these journals which traditionally have carried content analysis studies, namely *Public Opinion Quarterly, Journalism Quarterly, Journal of Abnormal and Social Psychology* (all major sources), and the *American Journal of Sociology, American Sociological Review, Journal of Applied Psychology, Social Forces, The Annals of the American Academy of Political and Social Science, Psychological Bulletin, Psychiatry, Library Quarterly*, and others.

**Major Problems of Content Analysis**

Much has been said elsewhere in this book about the necessity for having a clear statement of the problem before one starts out to investigate it. In content analysis, the same rule applies. The method involves far too much work of a rigorous, fatiguing nature to be indulged in haphazardly for purposes of fishing for a problem and a solution. Immense amounts of time and effort can be wasted in content analysis by an improperly prepared investigator. As Berelson says, "There is nothing in the natural, in the obvious, or in the inexplicable that will make the task easier.\(^5\)"

Assuming then, that you have a good idea, expressed in an abstract, generalized fashion, what are the major problems involved in investigating it through content analysis?
Sampling

Numerous social research does one except the necessity for having a sound sampling scheme. Such a plan is especially important in content analysis. People are wary. And one of the first problems that the content analyst faces, in almost any study he undertakes, is an immense potential collection of words—either distributed in space on the pages of newspapers, books, or magazines, or in time on radio or television. The material cannot be coded in its entirety; sampling in some form is a necessity if the problem is to be considered at all.

The word potential was italicized above because one of the chief difficulties in sampling of course, is the availability of relevant materials. The messages of the mass media are ephemeral products of the modern world; they flash out upon the scene and then are gone. At the time this is being written, for example, there is no place in America where one can find a complete current collection of American daily newspapers, to say nothing of a collection with deep runs into the past. Librarians tend to save only the prestige papers of the country. Thus, as a rule, analysts who want to generalize to the American press, must plan their study in advance and subscribe to a sample of newspapers. If one were interested in generalizing to American newspapers, a sample could be drawn by going to the current Editor & Publisher yearbook and giving each newspaper listed there a number and then drawing numbers randomly from the universe so established. Since there are many more small dailies than large dailies in the nation, however, such a plan would not necessarily represent the content typically seen by readers of American newspapers, most of whom live in large urban areas and read the large-city newspapers. A sample designed to represent what is available to readers, therefore, would have to be weighted in such a way that papers of larger circulation had a greater chance of being selected. Such a sample is described in the International Press Institute Flow of the News study.

Other newspaper sampling methods are given in the references.

After tens of publications have been selected, issues must still be sampled in turn. This presents additional problems to the content analyst. It is well known, for example, that American daily newspapers vary in thickness during the various days of the week, due primarily to the flow of advertising which reflects the weekend buying habits of Americans, with grocery ads whistling into view during the Christmas season. The number of such advertisements varies widely as well. The solution is to sample the text from such issues selectively. This is often done by hand; it is a laborious task.

Content analysis is the "com
Roy Carter. In this
all the Monday issues
of papers. Then a
Monday papers from
papers from the "a
structural weekly" the
newspapers and advertise
sample the totality of
this is a most useful
Establishing the Unit

Before content an
established. This is
which is to receive a
theme or assertion, the
Words are often the
encompassing subject
word or the others es
ception is simply t
as in some of the re
assuming reliability
have used words as
Often the analysis in
individual words, how
habit of Americans. Hence, papers of Thursday are typically fat with grocery ads while Saturday papers are slim. Similarly, magazines typically have their biggest issues of the year just before and during the Christmas shopping season. Radio and television stations fill the same number of hours every day with programs, but content varies widely to audiences ebb and flow with the hours and with the seasons.

What all this implies, of course, is that the messages to be studied must be sampled in time according to a plan which takes into account the various systematic factors which may influence the occurrence and nature of the messages. A sample of newspapers which included only content from Saturday afternoon issues would be a poor sample for most analyses, for example.

A method of time sampling sometimes used in newspaper content analysis is the "composite week" technique of Robert Jones and R.G. Center. In this method a "calendar" is constructed showing all the Monday issues, the Tuesday issues, etc., in the total sample of papers. Then a subsample is taken by randomly choosing N Monday papers from the Monday column of the calendar, N Tuesday papers from the Tuesday column, and so on. In this manner "constructed weeks" can be fashioned which will represent the flow of news and advertising through the days of the week, and also randomly sample the total time period involved in the study. For some analyses this is a most useful method.

Establishing the Units of Analysis

Before content analysis can begin, the basic coding unit must be established. This is the smallest division or segment of the content which is to receive a score. Typical coding units are the word, the theme or assertion, the item, time units, and the character.17

Words are often easier coding units to work with than the more encompassing subdivisions of content. It is fairly easy to define a word so that others can identify it in all its variations. If the coding operation is simply to note the presence or absence of certain words (as in some of the readability formulas, for example), a high degree of coder reliability can usually be obtained. Some studies which have used words as basic coding units appear in the references.18

Often the analyst is interested in larger units of meaning than individual words, however, and he feels that he must move to a larger
unit of analysis. If the hypothesis has been expressed in a general enough manner, however, it is sometimes possible to test it using a variety of coding units. If the hypothesis can be tested through some count of individual words or the relationship of individual words to one another, such a method certainly deserves at least a partial test of its feasibility. If the hypothesis can be as firmly noted with a single coding unit as with a complex one, the investigator is certainly justified in taking the easier course.

The theme or assertion is one of the most used units of content analysis. A theme is defined as "a simple sentence...an assertion about a subject-matter." For example, one might be interested in the attitude of an advice columnist toward various family members. Thus fathers, mothers, sisters, brothers might be significant "subject-matters." The statement, "Your father, though handsome, is a selfish, intolerant man," contains at least three assertions: namely, "Your father is handsome," "Your father is selfish," and "Your father is intolerant." It is evident that, unless the permissible referents are restricted, assertion analysis can rapidly explode a relatively small body of content into a large amount of data. The content analyst should be aware of this danger before he starts to use the method, so that he can be certain that he is coding only the content (and only the statements within the content) that are relevant to his hypothesis.

The student may wonder why content analysis makes relatively little use of more traditional units such as the sentence or the paragraph. There is no inherent reason why these units should not be used in some types of analysis, and indeed they are occasionally used. However, it is ordinarily more difficult to put these units into simple categories, if the categories are at all complex. What if a division (favorable, neutral, or unfavorable) is to be assigned to a sentence, as a unit? Obviously the sentence may contain conflicting directional statements, joined by such connectives as "but," "however," or "nevertheless." Coders, faced by such conflicting statements, may not operate reliably in coding the sentence. One time they may decide that the overall weight of the sentence is favorable, the next time they may decide that the weight is neutral or even unfavorable. The same criticism applies even more strongly to paragraphs as units.

Why then, is the most popular unit in content analysis the entire item, story, article, or editorial? Does not the same argument apply?
CONTENT ANALYSIS

In some cases, certainly, when coders attempt to place an entire article in a single complex category. Attempts to write directions for performing such feats nearly always produce frightful blunders. When careful directions are not written, coders tend to approach the task subjectively, and their classifications are not reliable. The wise student will always look for and examine with care the coder reliability figures in studies which report that complex categories were applied to coding units as large as whole stories. Often he will find that reliabilities simply are not reported. Using the entire article as a unit is desirable, however, in a variety of other situations when the classification categories are broad. For example, if one is interested in measuring attention paid to various categories of news, frequently the story has unity enough to enable coders to place it reliably into a single descriptive news category, or as most into two categories.[12] Using the large unit is desirable also if great amounts of material are to be coded. Somewhat lower reliabilities can then be tolerated because enough cases will be collected so that true relationships will show up to spite of a considerable amount of random coding error. A coding bias, of course, will continue to appear in spite of the large number of units coded.

Space and time units of analysis, such as the column inch (newspaper), the page or printed page (magazines), or the minute (radio and television), were widely used in early descriptive content analyses and are still used today.[10] When noble categories are to be employed, time and space measurements differ in the same way that item units and grammatical units differ. Moreover, several researchers have found that space and time units (which are difficult to measure precisely) correlate highly with item units; thus, the simpler natural item coding is being more frequently employed today.[13]

A character, or person, or a class of persons is sometimes used as a coding unit in content analysis. All relevant information about the character is sifted out of the article or story and classified. For example, Benson and Patricia Bulten classified characters in American magazine fiction according to whether they were “American,” “Anglo-Saxons and Northerns,” or “Others.” Then they examined each character as to his path, his occupation, whether he was “approved” or not, etc. They found that, in the stories analyzed, the Americans, Anglo-Saxons, and Northerns were more often seen as “approved” characters and were more often interested in heart goals.
than were the "Others." Recently, Jack Schwartz studied the portrayal of teachers in motion pictures in such terms as age, marital status, romantic involvement, relationships with students, etc. One of his conclusions was that in the movies, "... a successful compromise for the educator exists ... outside the academic pale; he leaves the teaching profession or marries someone with less education." The character unit has often been used, as the foregoing examples show, to examine the extent to which the mass media perpetrate the stereotypes of our society.

The Context Unit

The basic coding unit is the smallest division of context to receive a score. Sometimes, however, a score cannot be given solely from an examination of the basic unit; the unit can be coded reliably only in terms of its context. Hence, a context unit—the largest division of context which may be consulted by a coder in order to assign a score to a basic coding unit—is often used in content analysis. Suppose, for example, that you were to categorize all assertions made by a source about Communists as being favorable, unfavorable, or neutral. Suppose, further, that you encountered the following assertion: "The Communists are taking over the world bit by bit." In isolation, it is impossible to code the assertion accurately. It sounds neutral or ambiguous. Only when it is placed in a context unit is the direction apparent. For example, if the sentence appears in a speech entitled, "The Red Menace," delivered at a National Association of Manufacturers Convention by a conservative Republican senator, then the correct code probably would be unfavorable. If the assertion appeared in a speech by a Russian astronaut to a world gathering of Communist youth, however, the correct code would undoubtedly be favorable. Some limits are usually placed on the size of the context unit which the coder is permitted to scan. Ultimately the unit could become so large that the coding of the basic unit would again become unreliable. A historian trying to analyze Jefferson's quotations concerning the press in the context of the president's entire life might encounter this problem. Then, too, as the context unit becomes larger, more time is spent scoring each coding unit. Usually, therefore, a compromise must be sought between the desirable and the possible.

A final word on the selection of units of analysis is in order.
It is certainly desirable that the student be familiar with the definitions, units, codes, and categories other researchers have used before him. He can save valuable time by following the disciplined measures others have devised. Yet he should not slavishly follow the past. From time to time, as he is developing his idea, he should attempt deliberately to break with the conventions and try to see his problem with fresh and clear vision. It is easier to say this, of course, than to do it. But it is often in such contemplative moments that insights are born which advance science not by inches but by miles.

**Testing Coder Reliability**

At an early stage in the research, and at intervals thereafter, the method being used in the analysis should face tests of its reliability. This means, in general terms, that the investigator puts different coders to work on the same content and checks to see whether they apply the method in the same way and obtain the same results. It may also mean that the same coders, after a time lapse, are required to recode some material coded earlier to see whether they will do it the same way the second time. Coder reliability tests are extremely important and should never be neglected unless the codes being used are so obvious that even a severe critic will concede that "on the face of it" they can be reliably used. In many cases the chief investigator or his main analysts gradually learn to make many subtle qualifications and restrictions in coding content. To the extent that these qualifications and restrictions are not plainly spelled out and incorporated in the coding directions they cannot be communicated to others, and the method is likely to become increasingly unrepeatable. A coder reliability check taken at intervals in the research, using newly trained coders, will forcibly remind the researcher that his intuitions and subtle feelings for content must ultimately find concrete verbal expression if others are to use his method. Reputable journals are gradually becoming reluctant to print analyses in which coder reliabilities have not been tested.

Various methods of making coder reliability tests are mentioned in the references; hence illustrations of only two methods will be given here. Suppose one purpose of the analysis is to estimate the average number of page one national/international stories per issue of a number of newspapers. We might decide to test coder reliability.
in this instance by taking fifty newspapers from the sample and having each coder go through the papers, counting the number of national-international stories on the front page of each. We then could make a correlation matrix giving the scores assigned to each newspaper by each pair of coders. In Figure 1, for example, each mark in the scattergram represents the scores given to a particular front page of a paper by Coder A (on the vertical axis) and Coder B (on the horizontal axis). Inspection shows a high degree of agreement between these two coders. The Pearsonian product-moment correlation between the two sets of scores is .936. We might go on, from this point, and figure the product-moment correlations between the scores given by Coder A and Coder 5, then Coder B and Coder C, and so on. Finally, we might present the average correlation as our estimate of the coder reliability in this study.

**FIGURE 1**

Scattergram showing the number of national-international stories found on the front pages of fifty newspapers by two coders.

<table>
<thead>
<tr>
<th>Coder A</th>
<th>15</th>
<th>13</th>
<th>11</th>
<th>9</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>14</td>
<td>12</td>
<td>10</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coder B</th>
<th>15</th>
<th>13</th>
<th>11</th>
<th>9</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>14</td>
<td>12</td>
<td>10</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

r = .916

**CONTENT ANALYSIS**

What if the coders are disagreeable for some reason; the process of scoring the stories but coding it is often helpful but to look into this involved are ordinary codes and the percentage of age of stories on who of stories found btested. Thus, in 32% of the national-inter found 21%, of those found by Coder 2, these two coders should be divided, which is 343.0 which coders are.

(In the example, the sentences long and.

In some texts, for example, a 5 as juvenile, adult, involved in finding; be merely the men divided by the but all possible pairs first since the ap analysis reports, 2 percentage of age coder reliability be coding conditions, to the fact that the

**Testing for Validity**

Validity may be the technique measure.
What if the correlation coefficient is low and the scores of the two coders are disappointingly different? Obviously, at least two opportunities for errors are present: (1) A coder may be missing stories in the process of scanning the papers; or (2) A coder may be finding stories but coding them carelessly for one reason or another. Thus, it is often helpful not only to check the total score given a paper but to look into the matter further and find out whether the coders involved are coding the same stories in the same way. Such a procedure often results in a statistical measure referred to as the average percentage of agreement, which is, in one definition, the number of stories on which two coders agree divided by the total number of stories found by both the coders, averaged over all pairs of coders tested. Thus, suppose that in the example given, Coder B found 242 national-international stories in the fifty papers and Coder A found 241. Of the 242 stories found by Coder B, 231 were also found by Coder A. What is the percentage of agreement between these two coders? The answer is 96 per cent, because the 241 stories should be divided by the total number of stories found by both coders which is 483. An analysis of this type helps the researcher find which coders are making mistakes and what those mistakes are. (In the example, Coder A was counting stories which were less than 2 inches long and Coder B was not.)

In some test coders are given a certain number of units to code (for example, a certain number of stories about crime, to be coded as juvenile, adult, or mixed crime). In this case, there is no error involved in finding the stories, and the percentage of agreement will be merely the number of items coded the same way by two coders divided by the total number of items in the test and averaged over all possible pairs of coders. This test obviously is easier than the first since the opportunity for error is less. In reading context analysis reports, it is important to note specifically how the reported percentage of agreement was obtained. Generally speaking, those coder reliability tests are most respected which are done under natural coding conditions, using natural materials, and without alerting coders to the fact that they are participating in a special test of their ability.

**Testing the Validity of Results**

Validity supposedly answers the question, "Does the measurement, technique measure what it is supposed to measure?" In the hope-