Sudan University of Science and Technology
Graduate College
College of Languages

"Content Analysis", An Effective Tool of Investigation in Social Sciences

By:
Kamal Hussein Maghazi

Abstract:

Content analysis, as a tool of inquiry, has become an authorized approach, especially for human study researches, as it analysis the "text in the context", and this is the principal factor that distinguishes this method from other approaches.

Content analysis uses mathematical and statistical formulas to measure out the collected data. This gives it accuracy and credibility as it quantifies reliability and validity for every investigated item.

Above all this, one of the important characteristics of this method is that it runs the required procedures far from any subjective effect of the analysts.

المستخلَص:

أضحى "تحليل المحتوى"، كأداة للتحليل، من الطرق والوسائل المعتمدة التي يُعتد بها، وعلى الخصوص في الدراسات الإنسانية، وذلك يأتي من اتباع هذه الطريقة مبدأ تحليل النص "سواءً أكان مكتوبًا، منطوقًا، مشاهَدًا، أو حتى كان عبارة عن إيماءة" وفق الظروف المحيطة به ساعة حدوثه. هذا هو العامل الرئيس الذي ميّز هذه الأداة عن غيرها من أدوات التحليل النقدي.

تستخدم وسيلة "تحليل المحتوى" معادلات رياضية وإحصائية لقياس البيانات، وهذا ما يكسب هذه الطريقة الدقة نظرًا لقياسها لموثوقية وصلاحية الموضوعات محل التحليل.

بالإضافة إلى ما سبق، فإن العمليات التحليلية في طريقة "تحليل المحتوى" تُجرى بعيدًا عن أية تأثيرات ذاتية أو تدخُّلات من جانب الذين يقومون برصد البيانات وتحليلها.

Introduction:

Content analysis is potentially one of the most important research techniques in the social sciences. The content analyst views data as representations not of physical events but of texts, images and expressions that are created to be seen, read, interpreted and acted on for their meanings, and must therefore be analyzed with such uses in mind. Analyzing texts in the contexts of their uses distinguishes content analysis from other methods of inquiry.

Methods in the natural sciences are not concerned with meanings, contents, intentions and reference. These scientists hardly reflect on their own conceptions of nature, excluding their conceptions from their object of study by dismissing them as subjective in contrast to what can be determined through detached observation and objective measurement. Where social researchers adopt natural scientific methods of inquiry, the epistemology that is inscribed in such methods prevents them from addressing what matters most in everyday social life: human communication, how people coordinate their lives, the commitments they make to each other and to the conceptions of society they aspire to, what they know, and why they act. Certainly, content analysis is not the only research method that takes meanings seriously, but it is a method that is both powerful and unobtrusive. It makes sense of what is mediated between people-textual matter, symbols, messages, information, mass-media content and technology supported social interactions-without perturbing or affecting those who handle that textual matter.

In the 1980s, content analysis was a research method that had entered the psychological and social sciences, but was used mainly in journalism and communication research. At that time, the amount of human effort required to collect, transcribe and code textual data made content analysis a time consuming and labor-intensive effort. Today, content analysis has become

an efficient alternative to public opinion research, a method of tracking markets, political leanings, and emerging ideas; it is used as a way to settle legal disputes and as an approach to the exploration of individual human minds- not to dwell on the many improvements that content analysts have made in traditional content analytic inquiries of the mass media. Despite remarkable progress, content analysts can hardly claim to have met the challenges of this new era. The imagined analytical potential is far ahead of what can be done today, fueling the work of many developers of new analytic tools.

Content analysis Defined

The term content analysis is about 60 years old. Webster's Dictionary of the English Language included the term in its 1961 edition, defining it as "analysis of the manifest and latent content of a body of communicated material (as a book or film) through classification, tabulation and evaluation of its key symbols and themes in order to ascertain its meaning and probable effect". The intellectual roots of content analysis, however, can be traced far back in human history, to the beginning of the conscious use of symbols and voice, especially writing. This conscious use, which replaced the magical use of language, has been shaped by the ancient disciplines of philosophy, rhetoric, and cryptography. It has also spawned religious inquisitions and political censorship on the part of ruling establishments. Today, symbolic phenomena are institutionalized in art, literature, education, and the mass media, including the Internet. Theoretical and analytical concerns are found in such academic disciplines as anthropology, linguistics, social psychology, sociology of knowledge, and the comparatively younger field of communication studies. Many practical pursuits have grown from these fields: psychotherapy, advertising, politics, the arts, and so on. Virtually all disciplines within the whole spectrum of the humanities and the social sciences, including those that seek to improve the political and social conditions of life, are concerned with the functions and effects of symbols, meanings, and messages. In recent years, the emergence of the information society has moved the minutiae of communication-texts, contexts, images,

interfaces, and, above all, information into the very center of researchers' attempts at self-understanding.

However ancient the roots of analyzing symbolic and textual matter might be, today's content analysis is significantly different, in aim and in method, from that of the past. Contemporary content analysis has three distinguishing characteristics:

First, content analysis is an empirically grounded method, exploratory in process, and predictive or inferential in intent. Many of our current concepts relating to language are of Greek origin; for example, the words sign, significance, symbol, and logic all have Greek roots. However, the ancient Greeks' interest in language was largely prescriptive and classificatory, not empirical. Aristotelian logic set the standards for clear expression, and much of rhetorical theory was directed toward a normative conception of persuasive argumentation. Science that explores rather than declares is a relatively recent accomplishment. Only a century ago, George Boole and his contemporaries believed that the brain works according to (Boolean) logic and that human conduct is entirely rational. However, computers built on this logic turned out to be rather disappointing thinking machines. Empirical research in psychology is replacing Aristotelian categories in favor of a "psycho-logic." And we no longer measure human communication against the ideal of transmitting information. Instead, we inquire into what happens to the relationships between people who converse with one another.

With new conceptualizations and an empirical orientation, contemporary content analysts join other researchers in seeking valid knowledge or practical support for actions and critique. However, unlike researchers who employ other empirical techniques, content analysts examine data, printed matter, images, or sounds-texts-in order to understand what they mean to people, what they enable or prevent, and what the information conveyed by them does. These are questions for which natural scientists have no answers and for which their methods are generally insensitive.

Second, contemporary content analysis transcends traditional notions of symbols, contents, and intents. This may be seen in the evolution of the concept of communication, in how the development of media technologies

has shaped our attention to communication, and in the role of culture in assigning significance to what is being analyzed. I would argue that in recent years our awareness of communication has undergone four conceptual revolutions, as described below, and probably is in the midst of a fifth:

- The idea of ***messages***: the early awareness not only that verbal discourse is movable when written, but that writing has predictable effects. This awareness emerged in ancient Greece when messengers were used as the carriers of significance, history became documented, laws of the land were laid down in writing, and written instructions built organizational structures, directed events, and influenced (and possibly deceived) their receivers or the public. The concept of a message was a precursor of the rhetorical exploration of language. Tropes, syllogisms, and meanings came to be thought of as inherent qualities of speeches, letters, or documents. But a message is the metaphorical container of all these, a "container of content, " a vehicle for shipping meanings from one place to another-for example, when we now leave a message for someone on an answering machine or say that a message was meaningful (full of meanings) or meaningless (void of meanings) .
- The idea of ***channels***: the awareness of the constraints that every medium imposes on human communication. This awareness came with the increased reliance on different media of communication and served to explain their limitations: The alphabet limits what one can say in writing; the telephone confines communication to sound; and a television station can air no more than what is transmittable without interference from other stations, appealing to large audiences, and deemed profitable by its sponsors. The channel metaphor conjures images of canals and pipes with restricted capacities for shipping messages (with their contents) of certain forms and volumes.

- The idea of ***communication***: the awareness of the relational space between senders and receivers, of the processes through which interpersonal relation are negotiated, social structures are constituted,

and members of large populations come to know about each other. This awareness developed as an offshoot of the growth in mass media. By producing and disseminating identical messages-news and entertainment-to everyone, the mass media promised to be an agent of sharing, of building community relationships, of democratization, ideally, worldwide. Modeling themselves on the idea of mass production, the mass media also made us aware of where this one way model failed: in interpersonal conversation, point-to-point telephone communication, public debate, and dialogue. In U.S. culture, mass-media technology has become synonymous with progress, and communication is understood as the cure for most social problems-for example, we often blame lack of communication or miscommunication when interpersonal as well as national conflicts arise.

- The idea of ***systems***: the awareness of global, dynamic, and technologically supported interdependencies. This idea emerged with the growth of communication networks-telephone nets, wire services, mass-media systems, and most recently the Internet-transforming commerce, politics, and interpersonal relationships, creating networks whose properties have so far defied attempts to theorize them adequately. Unlike the one-way mass media, systems are marked by the interactivity and simultaneity of parallel communication on a massive scale and with the potential of nearly universal participation.
- The idea of computation: the awareness of the algorithmic nature of certain routine cognitive and social processes and their implementation in increasingly powerful computers. The processing of digital data in place of cognitive and social practices, along with the ability to reproduce these data in visual and textual forms for reading, rearticulating, and disseminating by and to ideally everyone, is encouraging an entirely new literacy that undercuts traditional organizational structures, including national boundaries. The fluidity and enormous complexity that computation has introduced into almost all spheres of life amplify the possibilities for scientific exploration as well as present unprecedented challenges for collective understanding.

This rather sketchy history of communication suggests that researchers who are concerned with texts can no longer focus only on symbols or representations, nor can they limit themselves to questions about "who says what, through which channels, to whom, and with which effects" (Lasswell, 1960) . The popular and simplistic notion of “content" has outlived its explanatory capabilities as well: content, the what of a communication, an entity that authors think they enter into messages and ship to remote receivers, who remove it for what it is and henceforth share it among others. This bizarre notion leads to authors as authorities of what they put into messages and to the conception of content analysts as experts who provide objective accounts of what messages were intended to conveyor actually contain.

History of Content Analysis

Empirical inquiries into the meanings of communications date back to theological studies in the late 1600s, when the Church found the printing of nonreligious materials to be a threat to its authority. Such inquiries have since mushroomed, moving into numerous areas and becoming the backbone of communication research. This chapter discusses several stages in the history of content analysis: quantitative studies of the press; propaganda analysis during World War II; social scientific uses of the technique in studies of political symbols, historical documents, anthropological data, and psychotherapeutic exchanges; computer text analysis and the new media; and qualitative challenges to content analysis.

Content analysis entails a systematic reading of a body of texts, images, and symbolic matter, not necessary from an author's or user's perspective. Although the term content analysis did not appear in English until 1941 (Waples & Berelson, 1941 , p. 2; cited in Berelson & Lazarsfeld, 1948) , the systematic analysis of text can be traced back to inquisitorial pursuits by the Church in the 1 7th century. Religions have always been captivated by the written word, so it is not surprising that the first known dissertations about newspapers were defended in 1 690, 1 695, and 1 699 by individuals pursuing academic degrees in theology. After the advent of the printing press, the Church became worried about the spread of printed matter of a nonreligious nature, and so it dealt with newspaper content in moralizing

terms (Groth, 1948 , p. 26) . Surprisingly, in spite of the rhetorical tradition of ancient Greece, which was normative and oral in orientation, the 17th century contributed very little to the methodology of content analysis.

Probably the first well-documented quantitative analyses of printed matter occurred in 18th-century Sweden. According to Dovring's (1954 - 1955) account, these analyses were undertaken as the result of the publication of the Songs of Zion, a collection of 90 hymns of unknown authorship. The collection had passed the Royal Swedish censor, but soon after its publication it was blamed for undermining the orthodox clergy of the Swedish state church. When the collection became popular, it was said to be "contagious " and was accused of aiding a dissenting group. Outstanding in this case is the fact that literary scholars of good reputation participated in the controversy, which crystallized around the question of whether the songs harbored dangerous ideas and, if so, how. Scholars on one side made a list of the religious symbols in the songs and became alarmed. Those on the other side, however, found the very same symbols in established song books and so discounted the claimed difference. Then some scholars noted that the symbols in the songs occurred in different contexts and had acquired meanings that were different from those taught in the official church. A debate arose about whether the meanings should be interpreted literally or metaphorically. The interpretations came to be compared with the results of a German study of the outlawed Moravian Brethren, a religious sect whose members later immigrated to the United States. This process-of revising a method in response to criticism--continued until it became clear to both sides in the debate how the symbols in the Songs of *Zion* differed from the symbols used in the official songbooks and how this (in the end political) phenomenon could be explained. The controversy generated many ideas that are now part of content analysis and stimulated debates about methodology that continue today.

Early Content Analysis

The second phase in the intellectual growth of content analysis, which took place in the 1930s and 1940s, involved at least four factors:

- During the period following the 1929 economic crisis, numerous social and political problems emerged in the United States. Many Americans believed that the mass media were at least partially to blame for such problems as yellow journalism, rising crime rates, and the breakdown of cultural values.
- New and increasingly powerful electronic media of communication, first radio and later television, challenged the cultural hegemony of the newspapers. Researchers could not continue to treat these new media as extensions of newspapers, because they differed from the print media in important ways. For example, users of radio and television did not have to be able to read.
- Major political challenges to democracy were linked to the new mass media. For example, the rise of fascism was seen as nourished by the as-yet little-known properties of radio.
- Perhaps most important, this period saw the emergence of the behavioral and social sciences as well as increasing public acceptance of the theoretical propositions and empirical methods of inquiry associated with them.

In the 1930s, sociologists started to make extensive use of survey research and polling. The experience they gained in analyzing public opinion gave rise to the first serious consideration of methodological problems of content analysis, published by Woodward in a 1934 article titled "Quantitative Newspaper Analysis as a Technique of Opinion Research." From writings about public opinion, interest in social stereotypes (Lippmann, 1922) entered the analysis of communications in various forms. Questions of representations were raised, with researchers examining topics such as how Negroes were presented in the Philadelphia press (Simpson, 1 934); how U.S. textbooks described wars in which the United States had taken part, compared with textbooks published in countries that were former U.S. enemies (Walworth, 1938); and how nationalism was expressed in children's books published in the United States, Great Britain, and other European countries (Martin, 1936).

One of the most important concepts that emerged in psychology during this time was the concept of " attitude." It added evaluative dimensions to

content analysis, such as "pro-con" or " favorable-unfavorable, " that had escaped the rough subject matter categories of quantitative newspaper analysis. Attitude measures redefined journalistic standards of fairness and balance and opened the door to the systematic assessment of bias. Among the explicit standards developed, Janis and Fadner's (1943-1965) "coefficient of imbalance" deserves mention. Psychological experiments in rumor transmission led Allport and Faden to study newspaper content from an entirely new perspective. In their 1940 articles "The Psychology of Newspapers: Five Tentative Laws," they attempted to account for the changes that information undergoes as it travels through an institution and finally appears on the printed page.

The interest in political symbols added another feature to the analysis of public messages. McDiarmid (1937), for example, examined 30 U.S. presidential inaugural addresses for symbols of national identity, of historical significance, of government, and of fact and expectations. Most important, Lasswell (1938) , viewing public communications within his psychoanalytical theory o f politics, classified symbols into such categories as " self" and " others " and forms of "indulgence" and " deprivation. " His symbol analysis led to his "World Attention Survey, " in which he compared trends in the frequencies with which prestige newspapers in several countries used national symbols (Lasswell, 1941) .

Researchers in several disciplines examined the trends in scholarship, as reflected in the topics that representative journals published. Rainoff's (1929) Russian study regarding physics was probably the first of this kind, but the most thorough analyses were conducted in the field of sociology (Becker, 1930, 1932; Shanas, 1945) and later in journalism (Tannenbaum & Greenberg, 1961) .

Several factors influenced the transition from quantitative newspaper analysis, which was largely journalism driven, to content analysis:

- Eminent social scientists became involved in these debates and asked new kinds of questions.
- The concepts these social scientists developed were theoretically motivated, operationally defined, and fairly specific, and interest in

stereotypes, styles, symbols, values, and propaganda devices began to replace interest in subject matter categories.

- Analysts began to employ new statistical tools borrowed from other disciplines, especially from survey research but also from experimental psychology.

Propaganda Analysis

Berelson described content analysis a s the use of mass communications as data for testing scientific hypotheses and for evaluating journalistic practices. Yet the most important and large-scale challenge that content analysis faced came during World War II, when it was employed in efforts to extract information from propaganda. Before the war, researchers analyzed texts in order to identify “propagandists," to point fingers at individuals who were attempting to influence others through devious means. Fears concerning such influence had several origins.

Propaganda was used extensively during World War I (Lasswell, 1927), and the years between the two world wars witnessed the effective use of propaganda by antidemocratic demagogues in Europe. In addition, Americans tend to have deep-seated negative attitudes toward religious fanatics, and the lack of knowledge concerning what the extensive use of the new mass media (radio, film, and television) could do to people raised concerns as well. According to the Institute for Propaganda Analysis (1937) , propagandists reveal themselves through their use of tricks such as " name-calling, " employing " glittering generalities, " "plain folks" identifications, " card stacking, " " bandwagon " devices, and so on. Such devices could be identified easily in many religious and political speeches, even in academic lectures, and this approach to propaganda analysis led to a kind of witch-hunt for propagandists in the United States. Theories concerning subliminal messages, especially in advertising, raised widespread suspicion as well.

In the 1940s, as U.S. attention became increasingly devoted to the war effort, the identification of propagandists was no longer an issue. Nor were researchers particularly interested in revealing the power of the mass media of communication to mold public opinion; rather, military and political intelligence were needed. In this climate, two centers devoted to propaganda analysis emerged.

Harold D. Lasswell and his associates, having written on political symbolism, worked with the Experimental Division for the Study of Wartime Communications at the U.S. Library of Congress, and Hans Speier, who had organized a research project on totalitarian communication at the New School for Social Research in New York, assembled a research team at the Foreign Broadcast Intelligence Service of the U.S. Federal Communications Commission (FCC). The Library of Congress group focused on analyzing newspapers and wire services from abroad and addressed basic issues of sampling, measurement problems, and the reliability and validity of content categories, continuing the tradition of early quantitative analysis of mass communications (Lasswell, Leites, & Associates, 1965) .

The FCC group analyzed primarily domestic enemy broadcasts and surrounding conditions to understand and predict events within Nazi Germany and the other Axis countries, and to estimate the effects of Allied military actions on the war mood of enemy populations. The pressures of day-to-day reporting left the analysts little time to formalize their methods, and Berelson (1952) thus had little to say about the accomplishments of the FCC group. After the war, however, Alexander L. George worked through the volumes of reports that resulted from these wartime efforts to describe methods that had evolved in the process and to validate the inferences the researchers had made by comparing them with documentary evidence now available from Nazi archives. These efforts resulted in his book Propaganda Analysis (1959a), which made major contributions to the conceptualization of the aims and processes of content analysis.

The assumptions that propagandists are rational, in the sense that they follow their own propaganda theories in their choice of communications, and that the meanings of propagandists' communications may differ for different people reoriented the FCC analysts from a concept of "content as shared" (Berelson would later say " manifest") to conditions that could explain the motivations of particular communicators and the interests they might serve. The notion of "preparatory propaganda " became an especially useful key for the analysts in their effort to infer the intents of broadcasts with political content. In order to ensure popular support for planned military actions, the

Axis leaders had to inform; emotionally arouse, and otherwise prepare their countrymen and women to accept those actions; the FCC analysts discovered that they could learn a great deal about the enemy's intended actions by recognizing such preparatory efforts in the domestic press and broadcasts. They were able to predict several major military and political campaigns and to assess Nazi elites' perceptions of their situation, political changes within the Nazi governing group, and shifts in relations among Axis countries. Among the more outstanding predictions that British analysts were able to make was the date of deployment of German V weapons against Great Britain. The analysts monitored the speeches delivered by Nazi propagandist Joseph Goebbels and inferred from the content of those speeches what had interfered with the weapons' production and when. They then used this information to predict the launch date of the weapons, and their prediction was accurate within a few weeks.

Several lessons were learned from these applications of content analysis, including the following:

- Content is not inherent to communications. People typically differ in how they read texts. The intentions of the senders of broadcast messages may have little to do with how audience members hear those messages. Temporal orderings, individuals' needs and expectations, individuals' preferred discourses, and the social situations into which messages enter are all important in explaining what communications come to mean. Interpretations on which all communicators readily agree are rare, and such interpretations are usually relatively insignificant.

- Content analysts must predict or infer phenomena that they cannot observe directly. The inability to observe phenomena of interest tends to be the primary motivation for using content analysis. Whether the analyzed source has reasons to hide what the analyst desires to know (as in the case of an enemy during wartime or the case of someone needing to impress) or the phenomena of interest are inaccessible in principle (e.g., an individual's attitudes or state of mind, or historical events) or just plain difficult to assess otherwise (such as what certain

mass-media audiences could learn from watching TV) , the analyst seeks answers to questions that go outside a text. To be sure, the questions that a content analyst seeks to answer are the analyst's questions, and as such they are potentially at odds. With whether others could answer them and how. Quantitative newspaper analysts made inferences without acknowledging their own conceptual contributions to what they thought they found but actually inferred. Content is not the whole issue; rather, the issue is what can be legitimately inferred from available texts.

- In order to interpret given texts or make sense of the messages intercepted or gathered, content analysts need elaborate models of the systems in which those communications occur (or occurred) . The propaganda analysts working during World War II constructed such models more or less explicitly. Whereas earlier content analysts had viewed mass-produced messages as inherently meaningful and analyzable unit by unit, the propaganda analysts succeeded only when they viewed the messages they analyzed in the context of the lives of the diverse people presumed to use those messages.
- For analysts seeking specific political information, quantitative indicators are extremely insensitive and shallow. Even where large amounts of quantitative data are available, as required for statistical analyses, these tend not to lead to the “most obvious" conclusions that political experts would draw from qualitative interpretations of textual data. Qualitative analyses can be systematic, reliable, and valid as well.

Convinced that content analysis does not need to be inferior to unsystematic explorations of communications, numerous writers in the postwar years, such as Kracauer (1947, 1952-1 953) and George (1959a), challenged content analysts' simplistic reliance on counting qualitative data. Smythe (1954) called this reliance on counting an " immaturity of science" in which obj ectivity is confusedwith quantification. However, the proponents of the quantitative approach largely ignored the criticism. In his 1 949 essay "Why Be Quantitative?” Lasswell (1949/1965b) continued to insist on the quantification of symbols as the sole basis of scientific insights.

His approach to propaganda analysis produced several working papers but very few tangible results compared with the work of the FCC group of scholars. Today, quantification continues, although perhaps no longer exclusively.

Content Analysis Generalized

After World War II, and perhaps as the result of the first integrated picture of content analysis provided by Berelson (1952), the use of content analysis spread to numerous disciplines. This is not to say that content analysis emigrated from mass communication. In fact, the very "massiveness" of available communications continued to attract scholars who looked at the mass media from new perspectives. For example, Lasswell (1941) realized his earlier idea of a "world attention survey" in a large-scale study of political symbols in French, German, British, Russian, and u.s. elite press editorials and key policy speeches. He wanted to test the hypothesis that a "world revolution" had been in steady progress for some time (Lasswell, Lerner, & Pool, 1952) . Gerbner and his colleagues pursued Gerbner's (1969) proposal to develop "cultural indicators " by analyzing, for almost two decades, one week of fictional television programming per year, mainly to establish "violence profiles" for different networks, to trace trends, and to see how various groups (such as women, children, and the aged) were portrayed on u.s. television (see, e.g., Gerbner, Gross, Signorielli, Morgan, & Jackson-Beeck, 1979) .

Psychologists began to use content analysis in four primary areas. The first was the inference of motivational, mental, or personality characteristics through the analysis of verbal records. This application started with All port's (1942) treatise on the use of personal documents, Baldwin's (1942) application of "personal structure analysis" to cognitive structure, and White's (1947) value studies. These studies legitimated the use of written material, personal documents, and individual accounts of observed phenomena as an addition to the then-dominant experimental methods. A second application was the use of verbal data gathered in the form of answers to open-ended interview questions, focus group conversations, and verbal responses to various tests, including the construction of Thematic

Apperception Test (TAT) stories. In the context of TAT stories, content analysis acquired the status of a supplementary technique. As such, it allowed researchers to utilize data that they could gather without imposing too much structure on subjects and to validate findings they had obtained through different techniques. Psychological researchers' third application of content analysis concerned processes of communication in which content is an integral part. For example, in his "interaction process analysis" of small group behavior, Bales (1 950) used verbal exchanges as data through which t o examine group processes. The fourth application took the form of the generalization of measures of meaning over a wide range of situations and cultures (which derived from individualist notions of meaning or content). Osgood (1974a, 1974b) and his students found numerous applications for Osgood, Suci, and Tannenbaum's (1957) semantic differential scales and conducted worldwide comparisons of cultural commonalities and differences.

Anthropologists, who started using content analysis techniques in their studies of myths, folktales, and riddles, have made many contributions to content analysis, including the componential analysis of kinship terminology (Goodenough,1 972) . Ethnography emerged in anthropology, and although ethnographers often interact with their informants in ways that content analysts cannot interact with authors or readers, after ethnographers gather their field notes they start to rely heavily on methods that are similar to those that content analysts use.

Historians are naturally inclined to look for systematic ways to analyze historical documents, and they soon embraced content analysis as a suitable technique, especially where data are numerous and statistical accounts seem helpful. Social scientists also recognized the usefulness of educational materials, which had long been the focus of research. Such materials are a rich source of data on processes of reading (Flesch, 1948, 1951) as well as on a society's larger political, attitudinal, and value trends. In addition, literary scholars began to apply the newly available techniques of content analysis to the problem of identifying the authors of unsigned documents.

On the one hand, this proliferation of the use of content analysis across disciplines resulted in a loss of focus: Everything seemed to be content

analyzable, and every analysis of symbolic phenomena became a content analysis. On the other hand, this trend also broadened the scope of the technique to embrace what may well be the essence of human behavior: talk, conversation, and mediated communication.

In 1955, responding to increasing interest in the subject, the Social Science Research Council's Committee on Linguistics and Psychology sponsored a conference on content analysis. The participants came from such disciplines as psychology, political science, literature, history, anthropology, and linguistics. Their contributions to the conference were published in a volume titled Trends in Content Analysis, edited by Ithiel de Sola Pool (1959a). Despite obvious divergence among the contributors in their interests and approaches, Pool (1959a, p. 2) observed, there was considerable and often surprising convergence among them in two areas: They exhibited (a) a shift from analyzing the "content" of communications to drawing inferences about the antecedent conditions of communications and (b) an accompanying shift from measuring volumes of subject matter to counting simple frequencies of symbols, and then to relying on contingencies (co-occurrences) .

Data Analysis

Content analysis is a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use.

As a technique, content analysis involves specialized procedures. It is learnable and divorceable from the personal authority of the researcher. As a research technique, content analysis provides new insights, increases a researcher's understanding of particular phenomena, or informs practical actions. Content analysis is a scientific tool.

Techniques are expected to be reliable. More specifically, research techniques should result in findings that are replicable. That is, researchers working at different points in time and perhaps under different circumstances should get the same results when applying the same technique to the same data. Reliability is the most important form of reliability.

Scientific research must also yield valid results, in the sense that the research effort is open for careful scrutiny and the resulting claims can be upheld in the face of independently available evidence. The methodological requirements of reliability and validity are not unique to but make particular demands on content analysis.

The reference to text in the above definition is not intended to restrict content analysis to written material. The phrase "or other meaningful matter" is included in parentheses to indicate that in content analysis works of an, images, maps, sounds, signs, symbols, and even numerical records may be included as data that is, they may be considered as texts-provided they speak to someone about phenomena outside of what can be sensed or observed. The crucial distinction between text and what other research methods take as their starting point is that a text means something to someone, it is produced by someone to have meanings for someone else, and these meanings therefore must not be ignored and must not violate why the text exists in the first place. Text-the reading of text, the use of text within a social context, and the analysis of text-serves as a convenient metaphor in content analysis.

In the content analysis literature, scholars have provided essentially three kinds of definitions of this research method:

1. Definitions that takes content to be inherent in a text
2. Definitions that takes content to be a property of the source of a text
3. Definitions that takes content to emerge in the process of a researcher analyzing a text relative to a particular context.

Each of these kinds of definitions leads to a particular way of conceptualizing content and, consequently, of proceeding with an analysis. Berelson's original definition of content analysis is an example of the first kind. Berelson (1952) defined content analysis as "a research technique for' the objective, systematic and quantitative description of the manifest content of communication " (p. 18) . His requirement that content analysis be “objective “and "systematic " is subsumed under the dual requirements of replicability and validity in our definition. For a process to be replicable, it must be governed by rules that are explicitly stated and applied equally to all units of analysis. Berelson argued for " systematicity" in order to combat the human tendency to read textual material selectively, in support of

expectations rather than against them. Our requirement of validity goes further, demanding that the researcher's processes of sampling, reading, and analyzing messages ultimately satisfy external criteria. Replicability is measurable and validity is testable, but objectivity is neither.

Our definition of content analysis omits three of Berelson's further requirements. One is his insistence that content analysis be " quantitative. " Although quantification is important in many scientific endeavors, qualitative methods have proven successful as well, particularly in political analyses of foreign propaganda, in psychotherapeutic assessments, in ethnographic research, in discourse analysis, and, oddly enough, in computer text analysis. The ability of computers to crunch words as well as numbers is well-known. When a computer program is used to analyze words, the algorithms that determine the program's operation must embody some kind of theory of how humans read texts, rearticulate texts, or justify actions informed by the reading of texts. Reading is fundamentally a qualitative process, even when it results in numerical accounts. By including the attribute “manifest" in his definition, Berelson intended to ensure that the coding of content analysis data be reliable; this requirement literally excludes "reading between the lines, " which is what experts do, often with remarkable intersubjective agreement (I will have more to say on this topic later in this chapter) .

My chief objection to Berelson's definition, and numerous derivatives of that definition, is related to his phrase " description of the manifest content of communication. " It implies that content is contained in messages, waiting to be separated from its form and described. Berelson felt no need to elaborate on the crucial concept of "content" in his definition because for him and his contemporaries, at the time he was writing, there seemed to be no doubt about the nature of content-it was believed to reside inside a text.

Berelson's operationalization of the attribute "manifest" is telling. If sources, receivers, and content analysts have different interpretations of the same message, which is quite natural, Berelson's definition restricts content to what is common to all of these accounts, what everyone can agree to.

Gerbner (1985) starts from a similar assumption when he insists that mass-media messages carry the imprint of their industrial producers. For

him, too, content is right there to be described for what it is. However, Gerbner goes beyond Berelson's notion by suggesting that the messages of the mass media are revealed in statistical accounts of their contents. Mass-media audiences, he suggests, are affected by certain statistical properties of mass-produced messages of which neither mass producers nor mass audiences are conscious. This privileges content analysts' accounts over the readings by audience members. Shapiro and Markoff's (1997) definition equates content analysis with scientific measurement as well, specifically, with " any systematic reduction . . . of text (or other symbols) to a standard set of statistically manipulable symbols representing the presence, the intensity, or the frequency of some characteristics relevant to social science" (p. 14) . Its implicit representationalism is common in several definitions of content analysis. For example, in a recent textbook, Riffe, Lacy, and Fico (1998) start with the proposition that content is central to communication research but then assert that the purpose of content analysis is to describe " it" so as to make "it" amenable to correlations with other (noncontent) variables-as if content were a variable or thing inherent to mass-media messages. These examples demonstrate that the container metaphor for meaning still abounds in much of the communication research literature (Krippendorff, 1993) . The use of this metaphor entails the belief that messages are containers of meaning, usually one meaning per message, and justifies calling any analysis of any conventionally meaningful matter a content analysis, regardless of whether it counts words or offers in-depth interpretations. Clearly, this is an insufficient way to define content analysis.

Definitions of the second kind distinguished above tie the content analysis of texts to inferences about the states or properties of the sources of the analyzed texts (Krippendorff, 1 969a, p. 70; Osgood 1959, p. 35) . Shapiro and Markoff (1997), among others, have criticized such definitions as too limiting. Holsti(1969, p. 2 5) elaborates on this idea by committing content analysis to an encoding/decoding paradigm in which message sources are causally linked to recipients through encoding processes, channels, messages, and decoding processes. Holsti wants the content analyst to describe the characteristics of communications in terms of "what, " " how, " and " to whom" in order to infer their antecedents in terms of

"who" and " why" and their consequences in terms of "with what effects. " The last of these could be determined more directly if sources and recipients were accessible to observation or were able to inform the analyst honestly. When antecedents and consequences are not accessible to direct observation, the analyst must make inferences. I am sympathetic to Holsti's logic, but putting sources-senders and/or receivers-in charge of the validity of the inferences may not be the best way for the content analyst to capture all of the communicators' intents. Moreover, describing message characteristics in terms of " what, " "how, " and " to whom" fails to acknowledge the analyst's own conceptual contributions to what constitutes the appropriate reading of the analyzed texts and the relevance of this reading to a given research question.

The analyst's conceptual contributions to the reading of a text are specifically recognized in an approach called ethnographic content analysis (Altheide, 1987) ; unfortunately, however, this approach has not been clearly defined. Proponents of ethnographic content analysis oppose the sequential nature of traditional content analysis, suggesting instead that analysts be flexible in taking into account new concepts that emerge during their involvement with texts. This approach acknowledges the theory-driven nature of content analysis but also demands that the analytical process be closely linked to the communicators studied. Ethnographic content analysis is emic rather than etic in intent; that is, it attempts to rely on indigenous conceptions rather than on analysts' theory imposed conceptions. Although the preference for communicators' conceptions would appear to tie ethnographic content analysis to the second kind of definition noted above, by urging researchers to reflect on their involvement in the process, the approach acknowledges the possibility that researchers' theories can play a role in how analysis proceeds. The latter ties it more closely to the third kind of definition of content analysis.

Conclusion:

To conclude this paper, it should be said that the method of content analysis is the ideal approach to analyze texts as it looks to it in its context. And also for investigating all the images of the tested item. Results of the critical

content analysis are always reliable and valid and that comes from the analytic statistical tools it uses.

References:

- Berelson. (1952). Content Analysis. Longman. U.S.A.
- Bureau of International Information Programs. Department, US State. (2008) The Civil Rights Movement. USA.
- Fairclough, N. (2001) Language and Power (second edition). Harlow: Longman.
- Gee, J.P. (1990) Social Linguistics and Literacies: Ideology in Discourse. London: The Falmer Press.
- George, Yule. (2014). The Study of Language. Cambridge University Press.
- Hughes, Langston and Miltzer, Milton. (1990) African American History. USA: Scholastic, Inc.
- Jorgensen, Marianne & Phillips, Louise (2002) Discourse analysis as theory and method. London: Longman.
- Krippendorff, J.S. (1933). Research Literature and Methods. Collins & Harper, U.S.A.