

# Six Questions for Every Content Analysis: Revisiting the Stemler's Overview

Usman Madugu<sup>1</sup>, Fatai Adesina Ibrahim<sup>2</sup>

<sup>1</sup>Department of Public Administration, Faculty of Management Science University of Abuja, Abuja, Nigeria

<sup>2</sup>Department of Public Administration, Federal University Lokoja, Nigeria

DOI: <https://dx.doi.org/10.47772/IJRISS.2024.8120233>

Received: 30 December 2024; Accepted: 03 January 2025; Published: 15 January 2025

## ABSTRACT

Content analysis (CA) is a technique widely used in various fields to systematically and objectively interpret textual, visual, or multimedia data. It facilitates the categorisation of phenomena or events for enhanced analysis and interpretation. In other words, CA is a technique for investigating social reality, which involves inferring attributes of a non-manifest context from attributes of a manifest text. This paper revisited Stemler's work, "an overview of content analysis" with specific emphasis on salient questions every content analysis effort ought to address. Following the review of the Stemler's, it is found that addressing the indispensable questions divulges that not all documents are amenable to CA under three conditions; substantial portions missing, improper recording and uncodability of document(s) owing to ambiguity of content or missing passages. This study also revealed some fatal flaws that inhibit the efficacy of CA: over-reliance on word-counts to attribute relevance, faulty definitions of categories and non-mutually exclusive and exhaustive categories. Thus, it is recommended that every CA should be tailored-made to answer the six indispensable questions so as to increase its credence and utility, which in turn addresses theoretical underpinnings, methodological rigor, ethical considerations, and the dynamic interplay between content, context, and interpretation.

**Key words:** Content Analysis, Coding, Emergent coding, A priori coding, Reliability, Validity

## INTRODUCTION

Content analysis (CA) is a systematic and replicable method for condensing extensive material into fewer content categories according to clear coding principles (Berelson, 1952; GAO, 1996; Krippendorff, 1980; Weber, 1990). In the same vein, Holsti (1969) provides a comprehensive definition of CA as "any technique for making inferences by objectively and systematically identifying specified characteristics of messages" (p. 14). According to Holsti's definition, CA is not confined to textual analysis; it can also be utilised in other domains, such as coding student drawings (Wheelock, Haney, & Bebell, 2000) or analysing actions recorded in videotaped studies (Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). To facilitate replication, the approach must be applied exclusively to data that possess durability. CA allows researchers to systematically examine extensive data sets with relative efficiency (GAO, 1996). This technique can effectively facilitate the identification and articulation of the focal points of individual, group, institutional, or social attention (Weber, 1990). It also permits inferences to be drawn, which can subsequently be validated by alternative data collection methods. Krippendorff (1980) observes that "much CA research is driven by the quest for methods to deduce from symbolic data what would be prohibitively expensive, no longer feasible, or excessively intrusive through alternative techniques" (p. 51).

CA serves as an effective instrument for establishing authorship. One method for establishing authorship involves assembling a list of potential writers, analysing their previous works, and correlating the frequency of nouns or function words to substantiate the likelihood of each individual's authorship of the relevant material. Mosteller and Wallace (1964) employed Bayesian methods grounded in word frequency analysis to establish

Madison as the author of the Federalist Papers; more recently, Foster (1996) adopted a comprehensive strategy to ascertain the identity of the anonymous author of the 1992 book *Primary Colours*. CA is effective for analysing trends and patterns within documents. Stemler and Bebell (1998) conducted a CA of school mission statements to determine the primary purposes for why schools operate. A key study inquiry was whether the metrics employed to assess program efficacy (e.g., academic test scores) were congruent with the overarching program objectives or purpose. Moreover, CA offers an empirical foundation for tracking changes in public opinion/sentiment. Data gathered from the mission statements project in the late 1990s can be objectively compared to data acquired in the future to ascertain whether policy changes connected to standards-based reform have been reflected in school mission statements.

The objective of this study is to delineate the concept of CA while emphasizing on the salient questions every CA effort ought to address and their influence on research. The data for this paper were acquired through reviewing extant literature on CA while revisiting Stemler's work, "an overview of content analysis". In addition, this study aims to provide researchers with a structured yet flexible framework to interrogate implicit biases, amplify marginalized perspectives, and uncover latent meanings embedded in diverse forms of content. Hence, this paper equips scholars and practitioners with the tools to effectively design and deploy CA studies, fostering more nuanced understanding and impactful outcomes. This approach not only deepens the analytical process but also encourages reflexivity and accountability in scholarly inquiry.

### Conducting a Content Analysis

Conducting a content analysis (CA) involves systematically examining and interpreting content in various forms (text, images, videos, etc.) to identify patterns, themes, meanings or trends. This method is widely used in qualitative research, media studies, and social sciences. Hence, the precursors to conducting a critical CA lie in the capacity of researchers to provide answers to indispensable questions raised by Krippendorff (1980). In other words, Krippendorff (1980) posits that six key questions must be addressed in every CA to ensure the methodological rigor and validity of the research. The indispensable questions are:

- 1) Which data are analysed?
- 2) How are the data defined?
- 3) What is the population from which the data are drawn?
- 4) What is the context relative to which the data are analysed?
- 5) What are the boundaries of the analysis?
- 6) What is the target of the inferences?

1. **Which data are analysed?** This question identifies the specific set of texts, messages, or content that will be examined in the study.
2. **How are the data defined?** This focuses on the criteria used to delimit or define the data. For example, what constitutes a "unit of analysis" (e.g., a sentence, paragraph, or entire document)? Defining the unit of analysis is central to content analysis. Stemler (2001) emphasized that units can range from words, phrases, or themes to larger constructs like paragraphs or entire documents. For example, in social media studies, units might be individual tweets or user comments (Neuendorf, 2017).
3. **What is the population from which the data are drawn?** This specifies the larger universe or population that the data represent. For instance, are the texts drawn from newspapers, social media posts, or speeches?
4. **What is the context relative to which the data are analysed?** This involves the situational or cultural context that influences how the data are interpreted. Context helps give meaning to the content being analysed.
5. **What are the boundaries of the analysis?** This defines the scope and limits of the study, including its objectives, timeframe, and thematic focus.

6. **What is the target of the inferences?** This addresses what conclusions or insights the researcher aims to derive from the analysis, such as trends, patterns, or meanings.

## Steps Involved in Conducting Content Analysis

CA is a systematic research technique deployed to interpret and analyse textual, visual, or auditory data to identify patterns, themes, or trends. It is widely used in qualitative research, media studies, communication studies, and social sciences. Here is a structured guide to conducting a CA:

1. *Define Your Research Question and Objective:* This stage clearly articulates the research problem or question, and set the aims of the CA. In other words, clearly outline what you aim to achieve, i.e. what are you aiming to learn or understand through this analysis? Example: "How is climate change portrayed in online news articles?"
2. *Determine the Scope:* This step determines the source and scope of the material. This can be achieved in the following manner:
  - a. *Content Sources:* Decide where the content will come from (e.g., news articles, academic journals, books, websites, social media, or video transcripts).
  - b. *Time-frame:* Establish the period of data to be analysed (e.g., articles from 2020–2022).
  - c. *Sample Size:* Choose how many pieces of content to be analysed to ensure it is manageable yet representative.
3. *Select a Unit of Analysis:* This could be specific words, phrases, sentences, paragraphs, or entire documents. For instance: Analysing the frequency of specific keywords (e.g., "global warming," "sustainability").
4. *Develop a Coding Framework:* This involves choosing between *quantitative* (frequency of words or phrases) or *qualitative* (themes, narratives) approaches.
  - a. *Categories:* Define themes or variables to be analysed. For instance: Tone (positive, neutral, negative), topics, or recurring imagery. In other words, the use of a coding system to categorize data.
  - b. *Coding Rules:* Create clear guidelines for coding to maintain consistency.
5. *Pilot the Coding Scheme:* This entails testing of the coding framework on a small subset of data to ensure clarity and reliability (i.e. refine categories). Hence, revise the framework as needed.
6. *Collect the Data and Organise:* Gather the content based on your defined scope and organize it systematically for easy access during analysis. In other words, ensure data is prepared for analysis, such as converting audio/video into text.
7. *Analyse the Data:* This is the process of applying your coding framework to identify and record relevant patterns or themes. In addition, software tools (e.g., NVivo, Atlas.ti, MAXQDA etc) can be deployed for large datasets, if needed.
8. *Interpret the Findings:* Look for patterns, relationships, or anomalies in the data. Then, relate or connect the findings back to your research question(s). Finally, compare your findings with existing literature or theories.
9. *Ensure Reliability and Validity:*
  - a. *Reliability:* Have multiple coders analyse the same data to check for consistency.
  - b. *Validity:* Cross-check interpretations against original content to ensure accuracy.
  - c. *Triangulation:* Use multiple data sources or methods to validate findings.
  - d. *Reflexivity:* Be aware of your biases and how they might affect the interpretation.
10. *Present the Results:* It entails presentation of the synopsis or summary of your findings in a report, chart, or graph. Then, highlight key themes and their implications for your research question(s). Finally, address the

limitations of the analysis and suggest areas for future research.

## Conditions for Unsuitability of CA

Three potential issues may arise during the gathering of documents for CA. First, when a significant quantity of documents from the population is absent, the CA must be discontinued. Second, unsuitable records (e.g., those that do not conform to the specified criteria for analysis) must be eliminated, but a log of the justifications should be maintained. Ultimately, several documents may fulfil the criteria for analysis yet remain uncodable due to absent sections or unclear content/ material (GAO, 1996).

## Analysing the Data

The prevalent conception in qualitative research is that CA merely entails conducting a word-frequency count. The premise is that the most often spoken terms correspond to the most significant concerns. Although this may hold in many instances, there are numerous counter-arguments to contemplate when employing basic word frequency counts to draw conclusions regarding significant issues.

It is important to note that synonyms may be employed for stylistic purposes within a paper, thereby causing scholars to undervalue the significance of a topic (Weber, 1990). Additionally, consider that each word may not equally reflect a category. Regrettably, there are no established weighting processes; so, researchers must acknowledge this constraint while utilising word counts. Moreover, Weber emphasises that not all issues present the same level of difficulty in being raised. In conducting word frequency analyses, it is essential to consider that certain terms may possess numerous meanings. The term "state" may refer to a governmental entity, a condition, or a verb denoting "to articulate."

A useful guideline in analysis is to employ word frequency counts to pinpoint words of possible significance, followed by a Key Word In Context (KWIC) search to assess the consistency of their usage. Most qualitative research tools (e.g., NUD\*IST, HyperRESEARCH) enables researchers to retrieve the phrase in which a specific term was utilised, allowing them to view the word within its context. This technique will enhance the validity of the inferences drawn from the data. Some software programs, such as the updated/revised *General Inquirer*, can include artificial intelligence algorithms (systems) that distinguish between identical words used in different contexts or same word with different meanings in different milieus (Rosenberg, Schnurr, & Oxman, 1990). A variety of software programs exist to assist in doing CA.

CA encompasses much more than mere word counts. The technique's richness and significance stem from its dependence on data coding and categorisation. The fundamentals of categorisation can be encapsulated in these quotations: "A category is a collection of words sharing analogous meanings or connotations" (Weber, 1990, p. 37). Categories must be mutually exclusive and comprehensive (GAO, 1996, p. 20). Mutually exclusive categories occur when no unit is between two data points, and each unit is represented by a single data point. The necessity for comprehensive categories is fulfilled when the data language encompasses all recording units without exception.

**Emergent versus a priori coding:** Two methodologies for coding data exist, each governed by distinct regulations. In *emergent coding*, categories are formed after preliminary assessment of the data. The procedures to be adhered to are delineated in Haney, Russell, Gulek, & Fierros (1998). In the first premise, two individuals separately evaluate the content and generate a compilation of features that constitute a checklist. Second, the researchers exchange observations and resolve any discrepancies identified in their preliminary checklists. Third, the researchers utilise a unified criteria for independent coding application. Fourth, the researchers assess the reliability of the coding, recommending a 95% agreement and a Cohen's kappa of 0.8. Should the reliability level be deemed unacceptable, the researchers go over the preceding stages? Upon establishing reliability, the coding is adopted on a large scale. The concluding phase involves a systematic quality control assessment.

*A priori coding* involves the establishment of categories prior to investigation, grounded in a theoretical framework. Professional peers concur on the classifications, and the coding is applied on the data. Revisions



are carried out as needed, and the categories are refined to optimise mutual exclusivity and exhaustiveness (Weber, 1990).

**Coding units:** Coding units are perceived in various ways. The first method is to delineate them physically according to their inherent or intuitive boundaries. For example, newspaper articles, letters, and poems possess inherent boundaries. The next method is to syntactically delineate recording units which involve utilising the distinctions deployed by the author; such as words, phrases, or paragraphs. The third method of defining them is using referential units. Referential units denote the representation of a unit. A document may reference George W. Bush as "President Bush," "the 43rd president of the United States," or "W." Referential units are advantageous for inferring attitudes, values, or preferences. A fourth approach to establishing coding units involves the utilisation of propositional units. Propositional units provide a complicated approach to creating coding units, as they deconstruct the text to analyse underlying assumptions. For instance, in the sentence, "Investors suffered another setback as the stock market persisted in its decline," we would analyse it as follows: The stock market has been underperforming recently/Investors have been incurring losses (Krippendorff, 1980).

CA basically employs three types of units: sample units, context units, and recording units.

*Sampling units* will differ based on the researcher's interpretative framework; they may consist of words, phrases, or paragraphs. The sampling unit in the mission statements project is the mission statement.

*Context units* need not be independent or separately describable. They may intersect and encompass numerous recording units. Context units impose physical constraints on the type of data being recorded. In the mission statements project, the contextual units are sentences. This decision is arbitrary, and the contextual unit may comprise paragraphs or complete statements of purpose.

In contrast, *recording units* are seldom delineated by physical boundaries. In the mission statements project, the recording units are comprising the concepts pertaining to the goal of school articulated in the mission statements (e.g., cultivate responsible citizens or enhance student self-worth). A sentence stating, "The mission of Al-Amin school is to enhance students' social skills, develop responsible citizens, and foster emotional growth," could be categorised into three distinct recording units, with each concept assigns to a single category (Krippendorff, 1980).

**Reliability:** Weber (1990) asserts: "To draw valid inferences from the text, it is essential that the classification procedure be reliable, meaning it must be consistent: Different individuals should categorise the same text identically" (p. 12). Weber observes that "reliability problems typically arise from the ambiguity of word meanings, category definitions, or other coding rules" (p. 15). It is essential to acknowledge that the individuals who devised the coding scheme have frequently engaged so intimately with the project that they have created both explicit and implicit interpretations of the coding. The evident outcome is that the reported reliability coefficient is artificially enhanced (Krippendorff, 1980). To prevent this, a crucial stage in CA is the formulation of a comprehensive set of explicit recording instructions. These directives enable external coders to undergo training until reliability standards are fulfilled.

Reliability can be articulated in the subsequent terms as follows:

*Stability, or intra-rater reliability:* Can the same coder achieve same outcomes repeatedly?

*Reproducibility, often known as inter-rater reliability:* Do coding schemes result in the same text being classified under the same category by several individuals?

A method to assess dependability/reliability is to evaluate the percentage of concordance among raters. This entails summing the cases coded identically by both raters and dividing by the total number of cases. The issue with a percentage agreement method is that it fails to consider that raters are anticipated to concur a certain proportion of the time only by chance (Cohen, 1960). To address this deficiency, reliability can be assessed using Cohen's Kappa, which approaches 1 when coding is wholesomely reliable and declines to 0 when there is no agreement but only by chance (Haney et al., 1998).

Cohen (1960) notes that there are three assumptions to attend to in using this measure. First, the units of analysis must be independent. For example, each mission statement that is coded is independent of all others. This assumption would be violated if in attempting to look at school mission statements, the same district level mission statement is coded for two different schools within the same district in the sample. Second, the categories of the nominal scale must be independent, mutually exclusive, and exhaustive. Suppose the goal of an analysis is to code the kinds of courses offered at a particular school. Now suppose that a coding scheme is devised that has five classification groups: mathematics, science, literature, biology, and calculus. The categories on the scale would no longer be independent or mutually exclusive because whenever a biology course is encountered it also would be coded as a science course. Similarly, a calculus would always be coded into two categories as well, calculus and mathematics. Finally, the five categories listed are not mutually exhaustive of all of the different types of courses that are probable to be provided at an educational institution. A foreign language course cannot be sufficiently categorized by any of the five classifications. The final assumption in the application of kappa is that the raters function independently. In other words, two evaluators should not collaborate to reach a consensus on the rating they would provide.

**Validity:** A methodology is consistently utilized to address a research question. Consequently, the validation of inferences drawn from a singular analytic method necessitates the utilization of diverse information sources. The researcher should endeavour to incorporate a validation study within the design, if feasible. In qualitative research, validation is achieved by triangulation. Triangulation enhances the credibility of findings by integrating diverse data sources, methodologies, researchers, or theoretical frameworks (Erlandson, Harris, Skipper, & Allen, 1993). In the mission statements project, the research question seeks to ascertain the goal of the school from the institution's standpoint. To cross-validate the findings from a content analysis, school administrators and hiring decision-makers could be interviewed regarding the significance attributed to the school's mission statement during the recruitment of prospective teachers, thereby assessing the degree to which a school's values are genuinely represented by its mission statements. An alternative method to validate the inferences would be to conduct a poll among students and teachers concerning the mission statement to assess their awareness of the school's objectives. An alternative approach would be to assess the extent to which the principles articulated in the mission statement are being enacted in the classrooms. Shapiro and Markoff (1997) contend that CA is credible and significant only in so far as its conclusions correlate with other metrics. Examining the correlation between average student performance on cognitive assessments and the focus on cognitive outcomes articulated in school mission statements might augment the validity of the results. For additional discourse concerning the legitimacy of content analysis, refer to Roberts (1997), Erlandson et al. (1993), and Denzin & Lincoln (1994).

## CONCLUSION

In conclusion, the data analysed in CA can vary widely, depending on the medium being studied and the research question(s). Researchers must define what data will be analysed, how they will collect it, and how it will be categorized to draw meaningful conclusions. Succinctly, CA serves as an effective data reduction methodology. The primary advantage is in its methodical, verifiable method for condensing extensive text into fewer topic groups according to stated coding standards. It possesses the appealing attributes of being inconspicuous and effective in managing substantial quantities of data. The method of CA encompasses much more than mere word frequency counts. The limitations of word counts have been examined, and strategies for expanding CA to improve its utility have been proposed. Two critical deficiencies that undermine the efficacy of a CA are imprecise definitions of categories and categories that are neither mutually exclusive nor exhaustive. In addition, Stemler's framework stresses that ethical lapses can undermine the integrity and applicability of CA findings. Therefore, it is recommended that every CA should be tailored-made to address the six indispensable questions so as to increase its credence and utility, which in turn addresses theoretical underpinnings, methodological rigor, ethical considerations, and the dynamic interplay between content, context, and interpretation. Hence, by systematically addressing these questions, researchers can ensure clarity, consistency, and reliability in their CA technique.

## REFERENCES

1. Berelson, B. (1952). *Content Analysis in Communication Research*. Glencoe, Ill: Free Press.

2. Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, pp. 37- 46.
3. Crocker, L., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. Orlando, FL: Harcourt Brace Jovanovich.
4. Denzin, N.K., & Lincoln, Y.S. (Eds.). (1994). *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage Publications.
5. Erlandson, D.A., Harris, E.L., Skipper, B.L., & Allen, S.D. (1993). *Doing Naturalistic Inquiry: A Guide to Methods*. Newbury Park, CA: Sage Publications.
6. Foster, D. (1996, February 26). Primary culprit. *New York*, 50- 57.
7. Haney, W., Russell, M., Gulek, C., and Fierros, E. (Jan-Feb, 1998). Drawing on education: Using student drawings to promote middle school improvement. *Schools in the Middle*, 7(3), 38- 43.
8. Holsti, O.R. (1969). *Content Analysis for the Social Sciences and Humanities*. Reading, MA: Addison-Wesley.
9. Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications.
10. Krippendorff, K. (1980). *Content Analysis: An Introduction to Its Methodology*. Newbury Park, CA: Sage.
11. Kvalseth, T. O. (1989). Note on Cohen's kappa. *Psychological reports*, 65, 223- 26.
12. Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, pp. 159-174.
13. Mosteller, F. and D.L. Wallace (1964). *Inference and Disputed Authorship: The Federalist*. Reading, Massachusetts: Addison-Wesley.
14. Neuendorf, K. A. (2017). *The Content Analysis Guidebook (2nd Edition)*. Sage Publications.
15. Nitko, A.J. (1983). *Educational Tests and Measurement: An Introduction*. New York, NY: Harcourt Brace Jovanovich.
16. Roberts, C.W. (Ed.) (1997). *Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts and Transcripts*. Mahwah, NJ: Lawrence Erlbaum Associates.
17. Rosenberg, S.D., Schnurr, P.P., & Oxman, T.E. (1990). Content analysis: A comparison of manual and computerized systems. *Journal of Personality Assessment*, 54 (1 & 2), 298- 310.
18. Shapiro, G., & Markoff, J. (1997). 'A Matter of Definition' in C.W. Roberts (Ed.). *Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts and Transcripts*. Mahwah, NJ: Lawrence Erlbaum Associates.
19. Stemler, S. (2001). An overview of content analysis. *Practical assessment, research, and evaluation*, 7(1).
20. Stemler, S., and Bebell, D. (1998). An Empirical Approach to Understanding and Analyzing the Mission Statements of Selected Educational Institutions. Paper presented at the annual meeting of the New England Educational Research Organization. Portsmouth, New Hampshire. Available: ERIC Doc No. ED 442 202.
21. Stigler, J.W., Gonzales, P., Kawanaka, T., Knoll, S. & Serrano, A. (1999). *The TIMSS Videotape Classroom Study: Methods and Findings from an Exploratory Research Project on Eighth-Grade Mathematics Instruction in Germany, Japan, and the United States*. U.S. Department of Education National Center for Educational Statistics: NCES 99-074. Washington, D.C.: Government Printing Office.
22. U.S. General Accounting Office (1996). *Content Analysis: A Methodology for Structuring and Analyzing Written Material*. GAO/PEMD-10.3.1. Washington, D.C. (This book can be ordered free from the GAO).
23. Weber, R. P. (1990). *Basic Content Analysis*, 2nd ed. Newbury Park, CA. Wheelock, A., Haney,
24. Wheelock, A., Haney, W., & Bebell, D. (2000). What can student drawings tell us about high-stakes testing in Massachusetts? *TCRecord.org*. Available: [http://www.tcrecord.org/Content .asp? Content ID=10634](http://www.tcrecord.org/Content.asp?ContentID=10634).