**Essay on Content Analysis**

**Marilyn Harbert - JOUR689R - Computational Text Analysis**

Content analysis is a research method used to systematically (and usually quantitatively) describe the content of a type of communication.[[1]](#footnote-0) This could involve counting times each person in a debate says specific words to understand which politician focuses more on a specific issue, or the sentiment of communications about a specific topic.[[2]](#footnote-1) Berleson writes that content analysis has three uses, understanding the “characteristics”, “causes”, and “consequences” of content.[[3]](#footnote-2) The range of uses that Berelson lists highlights the wide potential of content analysis and starts to hint at the many differences we see in implementation and use of content analysis. These differences can be seen in the types of content analysis Neuendorf elucidates in her book: rhetorical analysis; narrative; discourse; structuralist; interpretative; conversation; critical; and normative.[[4]](#footnote-3)

While content analysis can be applied in several different ways, it has a few general common steps. First the researcher should choose the unit of analysis.[[5]](#footnote-4) The unit of analysis is the level of the concept of interest which the researcher’s findings will be made about and the research questions will be asked about.[[6]](#footnote-5) In content analysis common units of analysis include words, themes, or phrases.[[7]](#footnote-6) It is especially important because it shapes what the findings can apply to, whether they can be generalized at the level of books, or words, is very different.

Neuendorf argues that content analysis starts abstractly with theorization, a review of the literature, and the generation of research questions and a hypothesis.[[8]](#footnote-7) The researcher should define the independent and dependent variables of interest in their research questions and hypotheses.[[9]](#footnote-8) The independent variables are the things the researcher believes might be causing a specific effect, seen in the dependent variables.[[10]](#footnote-9) These are important because like the unit of analysis they help define the scope of the research study.[[11]](#footnote-10) The data measured should then be an operationalization of the concepts of interest, that is they should be good real world stand ins for the concepts the researcher is hoping to study.[[12]](#footnote-11) If they are not good stand ins for the concepts of interest (like using income to measure education) then the study is going to fail before it starts.

Downe-Wamboldt suggests that the next step is to create the categories to organize the content into, and conduct a pretest of the categories and the codebook or coding schemes.[[13]](#footnote-12) The pretest allows the researcher to refine the categories, and ensure that the coding is consistent and validly measures the concepts of interest.[[14]](#footnote-13) Any changes would then be tested again, until they are satisfactory.[[15]](#footnote-14)

When writing the codebook for content analysis, it should explain how each variable in the study is measured, the definitions of the categories to code the content into, how coders determine whether content fits into a category, and what they write or mark down to indicate whether or not or how much a unit of data fits into a category.[[16]](#footnote-15)

Once the codebook or coding scheme is perfected, the researcher can code the remainder of the data.[[17]](#footnote-16) If done by humans, the coding should be conducted by at least 2 people, and intercoder reliability (level of agreement on the coding of the data) should be calculated between the coders to ensure the categories are reliably applied.[[18]](#footnote-17) Then, the results can be written up to be reported.[[19]](#footnote-18)

While content analysis is a rigorous, generalizable method of analysis, it has its limitations. Content analysis generally (with some exceptions to its implementation) tries to quantify qualitative data. In doing so, there is a generalization that can lose sight of some of the interesting findings that lie in the details or the outliers. It also takes the data out of its context, which can result in the loss of crucial information needed to gain a full understanding of the situation. Anthropologist Clifford Geertz describes a method called “thick description” in which social situations are described in deepest detail in order to understand them deeply, rather than broadly.[[20]](#footnote-19) He describes the example of the wink - a wink can be playful, flirty, practiced, or imitative just to name a few different ways a single wink could be different.[[21]](#footnote-20) Through description of all the context and details of a social act such as a wink (thick description), one can tell the different winks apart and understand the meaning meant to be conveyed.[[22]](#footnote-21) When you conduct quantitative content analysis, you lose that deeper meaning, and group all winks together.

**Bibliography**

Berelson, Bernard. "Content analysis in communication research." *Media studies: A reader* 200209 (2000).

Downe‐Wamboldt, Barbara. "Content analysis: method, applications, and issues." *Health care for women international* 13, no. 3 (1992): 313-321.

Geertz, Clifford. "Thick description: Toward an interpretive theory of culture." In *The cultural geography reader*, pp. 41-51. Routledge, 2008.

Maseda, Barbara. "17. Text as Data: Finding Stories in Text Collections." *Towards a Critical Data Practice* (2021): 116.

Neuendorf, Kimberly A. *The content analysis guidebook*. sage, 2017.

1. Berelson, Bernard. "Content analysis in communication research." *Media studies: A reader* 200209 (2000). [↑](#footnote-ref-0)
2. Maseda, Barbara. "17. Text as Data: Finding Stories in Text Collections." *Towards a Critical Data Practice* (2021): 116. [↑](#footnote-ref-1)
3. Berelson, Bernard. "Content analysis in communication research." *Media studies: A reader* 200209 (2000). [↑](#footnote-ref-2)
4. Neuendorf, Kimberly A. *The content analysis guidebook*. sage, 2017. [↑](#footnote-ref-3)
5. Downe‐Wamboldt, Barbara. "Content analysis: method, applications, and issues." *Health care for women international* 13, no. 3 (1992): 313-321. [↑](#footnote-ref-4)
6. Neuendorf [↑](#footnote-ref-5)
7. Downe‐Wamboldt, Barbara. "Content analysis: method, applications, and issues." *Health care for women international* 13, no. 3 (1992): 313-321. [↑](#footnote-ref-6)
8. Neuendorf, Kimberly A. *The content analysis guidebook*. sage, 2017. [↑](#footnote-ref-7)
9. ibid. [↑](#footnote-ref-8)
10. ibid. [↑](#footnote-ref-9)
11. ibid. [↑](#footnote-ref-10)
12. ibid. [↑](#footnote-ref-11)
13. Downe‐Wamboldt, Barbara. "Content analysis: method, applications, and issues." *Health care for women international* 13, no. 3 (1992): 313-321. [↑](#footnote-ref-12)
14. ibid. [↑](#footnote-ref-13)
15. ibid. [↑](#footnote-ref-14)
16. Neuendorf, Kimberly A. *The content analysis guidebook*. sage, 2017. [↑](#footnote-ref-15)
17. Downe‐Wamboldt, Barbara. "Content analysis: method, applications, and issues." *Health care for women international* 13, no. 3 (1992): 313-321. [↑](#footnote-ref-16)
18. Neuendorf, Kimberly A. *The content analysis guidebook*. sage, 2017. [↑](#footnote-ref-17)
19. ibid. [↑](#footnote-ref-18)
20. Geertz, Clifford. "Thick description: Toward an interpretive theory of culture." In *The cultural geography reader*, pp. 41-51. Routledge, 2008. [↑](#footnote-ref-19)
21. ibid. [↑](#footnote-ref-20)
22. Geertz, Clifford. "Thick description: Toward an interpretive theory of culture." In *The cultural geography reader*, pp. 41-51. Routledge, 2008. [↑](#footnote-ref-21)