



Qualitative Research in the Study of Leadership

Content Analysis in Leadership Research

Article information:

To cite this document: "Content Analysis in Leadership Research" *In Qualitative Research in the Study of Leadership*. Published online: 09 Jan 2017; 93-122.

Permanent link to this document:

<https://doi.org/10.1108/978-1-78560-651-920152006>

Downloaded on: 24 October 2017, At: 10:42 (PT)

References: this document contains references to 0 other documents.

To copy this document: permissions@emeraldinsight.com

The fulltext of this document has been downloaded 3 times since 2017*

Users who downloaded this article also downloaded:

(2016),"Case Studies in Leadership Research", , pp. 59-91 https://doi.org/10.1108/978-1-78560-651-920152004

(2016),"Qualitative Research as Method", , pp. 31-55 https://doi.org/10.1108/978-1-78560-651-920152003



Access to this document was granted through an Emerald subscription provided by emerald-srm:368933 []

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

4

CONTENT ANALYSIS IN LEADERSHIP RESEARCH

AN ILLUSTRATIVE LEADERSHIP CONTENT ANALYSIS STUDY

Content analysis was applied in this cross-cultural study to assess the impact of a company's cultural diversity management on the competitiveness of the firm since the increased globalization has made an imperative for managers to implement strategies of diversity to ensure a harmonious work environment in which employees from different cultures and of different cultural identities can harmoniously coexist. To achieve the purpose of this study, the authors (Sultana, Rashid, Mohiuddin, & Mazumber, 2013) analyzed 30 scientific articles published in international peer-reviewed journals that reported company efforts to improve the management of cultural diversity and thereby enhance the competitiveness of the organization.

Each of the members of the research team content analyzed 10 articles by identifying the nodes and free nodes hierarchically using both inductive and deductive coding. All authors organized their codes into two levels: first level codes which were those that directly reflected on the meaning units and second level which included a set of first level code. These analyses revealed three distinctive categories: cultural diversity management, conducive work environment, and competitive advantage. Each of these categories was comprised of a number of subcategories or codes. For example, in the conducive work environment category, the authors identified subcategories such as accepting differences, reconciling personal and professional lives, and learning intercultural competencies. In addition, leadership emerged as a major subcategory in the cultural diversity management category. The coding process was augmented by the use of

a computer assisted qualitative analysis program (CAQDAS), NVivo. At the end of the coding process, the authors summarized five key dimensions of cultural diversity management essential to gain companies' competitive advantage as demonstrating creativity, innovation, image and marketing, higher productivity, and work force competencies in a conducive environment. The authors highlighted the importance of the research questions and structured the analysis in such a way that the results answered those questions.

This study, in contrast to many content analyses summarized in [Table 4-4](#) which used sophisticated quantitative analyses, is representative of the more quantitatively-oriented approach to content analysis. The authors did not conduct a formal reliability assessment but simply noted that the researchers identified categories that were semantically similar. However, with regard to validity, they make references to the transferability of their results to theory, mainly [Hofstede's \(2001\)](#) seminal work on cultural differences.

INTRODUCTION

Defining Content Analysis

Content analysis has been defined as a family of procedures for studying the contents of written or transcribed texts which enables the researcher to include large amounts of textual information and systematically identify its properties such as frequencies of the most frequently used key words. According to [Krippendorff \(2012\)](#), content analysis allows researchers to make replicable and valid inferences from data to their context. Texts can be defined broadly as books, articles, cartoons, graffiti, newspaper headlines, historical documents, transcripts of interviews, and advertising but also includes nontext-based materials such as TV segments, photography, video tapes, and materials captured from the performing and visual arts or any other communicative medium. Textual data may be in verbal, print, or electronic form and may be obtained from narrative responses, open-ended survey questions, interviews, focus groups, observations, or print media such as articles, books, or manuals ([Kondracki & Wellman, 2002](#)). Some definitions include references to visual communication messages (e.g., [Elo & Kyngäs, 2008](#)).

Content analysis may be used with either qualitative or quantitative data and analyzed inductively or deductively. The inductive approach, that is, the process of moving from the specific to the general, means that particular instances of a speech are observed and then combined into a larger statement. It is used when there are no previous studies dealing with the phenomenon. The deductive approach, on the

other hand, moves from the general to the specific (Burns & Grove, 2005). Deductive content analysis is used to retest existing data in a new context, testing categories, concepts, models, or hypotheses. It is usually based on earlier work such as theories, models, mind maps, and literature reviews (Hsieh & Shannon, 2005).

Language has long been recognized as a critical aspect of the leadership process because it is through communications between leaders and their followers that leadership is ultimately enacted (Bligh & Kohles, 2014). The authors suggest that leadership researchers continually seek out and utilize textual materials from diverse sources such as emails, letters, annual reports, executive speeches, presidential addresses, company newsletters, websites, and blogs. As a research tool, content analysis is used to determine the presence of certain concepts and themes within a corpus of text. Classical content analysis in the quantitative tradition, as Bligh, Kohles, and Meindl (2004b) note, focuses squarely on this linguistic map of the relationship between leaders and followers, as well as the subsequent transformation of contextualized data into numeric output. The authors go on to say that as a research method, content analysis is ideally suited for researchers who seek to demonstrate reliability using a variety of correlation coefficients discussed later in this chapter.

The focus of this chapter is on textual materials; Chapter 9 examines the role of sounds and sights in qualitative research. According to Titscher, Meyer, Wodak, and Vetter (2000), content analysis is the “longest established method of text analysis among the set of empirical methods of social investigation” (p. 55). Content analysis has gained significance in the first half of the twentieth century resulting from the dramatic expansion of mass communication (Mayring, 2000). But even before that, different approaches to analysis and comparison of texts from a hermeneutic perspective (e.g., Bible interpretations) and early analysis of newspaper articles and graphological procedures can be seen as early precursors of content analysis.

Content analysis covers a wide range of approaches and techniques ranging from purely qualitative to highly quantitative. Initially researchers used content analysis as either a qualitative or quantitative method; later it was primarily used as a quantitative technique, with text data coded into implicit categories and then described using statistics. This approach is sometimes referred to as quantitative analysis of qualitative data. While content is undoubtedly a technique that is applied to qualitative data, it is not in itself a qualitative method. In many ways, its emphasis on quantification and on reliable and replicable coding rules exemplifies the fundamentals of quantitative research (Bryman, 2004). However, purely qualitative analyses, as illustrated by the opening vignette, are alternatives to quantitative approaches. Qualitative content analysis tries to overcome the shortcomings of quantitative

content analysis such as providing answers to how the categories were derived by applying a systematic, theory-guided approach to text analysis. Qualitative content analysis a systematic, rule guided text analysis that preserves some of the strengths of quantitative content analysis and widens them to adopt a variety of qualitative procedures.

DESIGNING A CONTENT ANALYSIS STUDY

Neuendorf (2002) describes the process of conducting a content analysis study as a series of steps depicted in Fig. 4-1. As depicted in Fig. 4-1, the content analyst first researches the topic of interest, examines existing theory that has a bearing on the research question, and then goes through a series of design decision such as the use of manual coding versus computerized coding (or a combination of both). The next step involves the identification of the corpus of text (sampling), coder training, and establishing intercoder reliability which are described later in this chapter. *Insch, Moore, and Murphy (1997)*, in an early review of content analysis in leadership research, add some additional components such as specifying the unit of analysis, data collection, and assessing construct validity. But regardless of the specific procedure employed, content analysis is typically presented as a series of well-defined steps that take the analyst from the initial identification of the text to be analyzed to the derivation research questions to the final interpretation of the data.

In Fig. 4-1, the preparation, organization, and results phases of a content analysis are represented capturing both inductive and deductive processes. However, the stepwise procedure presented here should not be interpreted as prescriptive since there are no systematic rules for collecting and analyzing the data keeping in mind that the key feature of all content analysis procedures is the reduction of much word into a smaller number of categories.

DATA COLLECTION AND ANALYSIS IN CONTENT ANALYSIS

Data collection begins with the selection of the corpus (body) of text to be analyzed and reading it a number of times. It also involves the development of an understanding of the context in which the text is placed as well as an analysis of the previous knowledge the researcher brings to the project. Since content analysis is not unambiguously a qualitative or quantitative method, it is amenable to both qualitative (e.g., *Boyatzis, 1998; Mayring, 2000, 2002*) and quantitative analyses such as *t*-tests, correlations, and loglinear models (*Krippendorff, 2012; Neuendorf, 2002*). Building a dictionary, using a stop word list, selecting key words, segmenting a

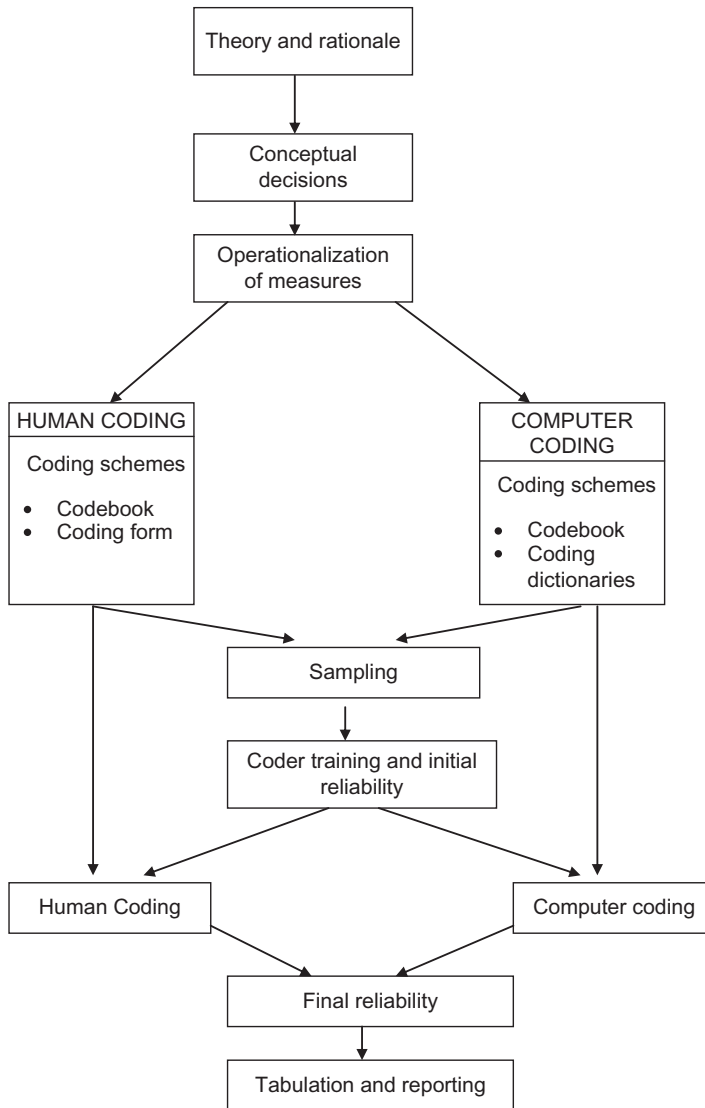


Figure 4-1 The process of content analysis research (Source: adapted from Neuendorf, 2002)

corpus of text into parts or phrases, collapsing codes into categories, using some type of CAQDAS software, are examples of issues that need to be decided on during the initial phase of the content analysis project. In addition, the content analyst needs to decide whether an entire set of texts will be analyzed or whether a sampling strategy will be used to select fewer, representative segments of text. The next decision before proceeding with the actual coding of a corpus of text concerns the

determination of the unit of analysis or recording units, which is the basic unit of text to be analyzed. Units of analysis can be words, sentences, paragraphs, or entire documents where the entire text is assigned to a category.

Data collection starts with reading all texts repeatedly to achieve immersion and obtain a sense of the whole. Then data are read word by word, line by line to derive the initial codes by highlighting the exact words from the text that appear to capture key words and concepts (Kuckartz, 2014). Next, the researcher approaches the text by taking notes of first impressions, thoughts, and initial analysis continuing this process until labels for codes emerge that are reflective of that one key thought. Content-bearing data must be converted into an acceptable format, such as numbers, categories, or codes. Coding is the heart and soul of text analysis. Codes often come directly from the text and then become the initial coding theme. A coding scheme includes the process and rules of data analysis that are systematic, logical, and scientific. The development of a good coding scheme is central to trustworthiness in research using content analysis.

Qualitative and Quantitative Approaches to Coding

One of the key elements in content analysis is the systematic coding of text. Krippendorff (2012) defines coding as the transcribing, recording, categorizing, or interpreting of given units of analysis (words, speeches, book, web pages, photographic images, etc.) into the terms of a data language so that they can be analyzed. Codes are building blocks for theory and model building and the foundation on which the analyst's arguments rest. Implicitly or explicitly, they embody the assumptions underlying the analysis. As Bernard and Ryan (2010) point out, a code can be used as an encryption, indexing, or measuring device (p. 87). Codes add information to the text (rather than reproducing the text) through the process of interpretation that simultaneously breaks down the text into meaningful chunks or segments. Thus, the coding process must include specific guidelines for defining the boundaries of the text associated with a particular code. How the text is segmented which is influenced by the data collection strategy and the way that strategy structures the resulting text.

Types of Coding

In content analysis, there are five major types of coding: theoretical coding, open coding, axial coding, selective coding, and thematic coding. Regardless of the types

of coding employed, researchers address the text regularly and repeatedly with the following basic questions (Strauss & Corbin, 1990a, 1990b, p. 77):

- What? Which phenomenon is mentioned?
- Who? Who are the persons involved? Which roles do they play? How do they interact?
- How? Which aspects of the phenomenon are mentioned (or not mentioned)?
- When? How long? Where? Time, course, and location
- How much? How strong? Aspects of intensity
- What for? With what intention, to which purpose?
- By which? Means, tactics, and strategies for reaching the goals.

Theoretical coding is the procedure for analyzing data which have been collected in order to develop a grounded theory. This procedure was introduced by Glaser and Strauss (1967) and further elaborated by Corbin and Strauss (2014). Coding here is understood as representing the operations by which data are broken down, conceptualized, and put back together in new ways. Theoretical coding is based on the constant comparison method or a constant going back and forth between data to develop theory returning to the collection of additional data and refining the emergent theory. During the whole process, impressions, associations, and questions are noted in code notes or memos. This process is known as *memoing* which is an integral process in most coding procedures (Flick, 2007). The content analysts writes a number of memos such as theoretical, reflective, analytical, or methodological memos which contain ideas, concepts, assumptions, or peculiarities in the text as they occur to the coder in the process of analysis.

Open coding is designed to express data in the form of initial concepts. For this purpose, data are segmented and expressions are classified by their units of meaning (single words, short sequences of words in order to attach concepts to them). Sometimes hundreds of codes results (Corbin & Strauss, 2014). In this step of the coding process, the initial codes are provisional, comparative, and grounded in the data (Charmaz, 2006, p. 37). The next step in the procedure is to categorize these codes by grouping them around phenomena associated with the data which are particularly relevant to the research question. The resulting categories are then linked to codes which are now more abstract than those used in the first step. Open coding may be applied by using various degrees of detail. A text can be coded line by line, sentence by sentence, or paragraph by paragraph, or a code can be linked to whole texts (an interview protocol, a case). The result of open coding should be a list of

the codes and categories that were attached to the text. This coding phase should be complemented by memos that are produced for explaining and defining the content of codes and categories.

Axial coding follows open coding as the next step to refine and differentiate the categories resulting from open coding. From the multitude of categories that were originated in open coding, those are selected that seem to be most promising for further elaboration. The axial categories are enriched by their fit with as many passages as possible. Most importantly, the relations between categories and their subcategories are clarified or established. Axial coding is summarized as follows:

Axial coding is the process of relating subcategories to a category. It is a complex process of inductive and deductive thinking involving several steps. These are accomplished, as with open coding, by making comparisons and asking questions. However, in axial coding the use of these procedures is more focused, and geared toward discovering and relating categories in terms of the paradigm model (Strauss & Corbin, 1990a, 1990b, p. 114).

Kuckartz (2014) added that axial coding focuses on one specific category which is examined through six classes to include “phenomena, causal conditions, context, intervening conditions, action strategies, and consequences” (p. 25).

Selective coding continues axial coding at yet a higher level of abstraction. The aim of this step is to elaborate the core category around which the other developed categories can be grouped and by which they are integrated. The result of this step may be one single category capturing one central phenomenon. Other coding schemes exist such as the one proposed by Saldaña (2009) who describes a coding cycle consisting of first cycle codes (corresponding to open coding) as well as a variety of specific coding techniques such as *in vivo* or holistic coding and second cycle coding which includes axial and theoretical coding in addition to a number of other techniques. The primary goal of the second coding cycle is “to develop a sense of categorical, thematic, conceptual, and/or theoretical organizations for the array first cycle codes” (Saldaña, 2009, p. 149).

Coding is terminated at the point of theoretical saturation when further coding, enrichment of categories, etc. no longer provide new knowledge. At that time, the procedure is flexible enough that the researcher can re-enter the same source text and same codes from open coding with a different research question and aim at developing and formulating a theory of a different issue. Coding, like factor analysis in quantitative analysis, is a data reduction technique that results in increasingly higher levels of abstraction.

Thematic coding was initially developed by Strauss (1987) and elaborated on by Boyatzis (1998). According to Boyatzis (1998), thematic coding is a multistep procedure that can be performed inductively or deductively. According to the author, a good code should have five elements:

1. A label
2. A definition of what the theme concerns
3. A description of how to know when the theme occurs (i.e., how to “flag” the theme)
4. A description of any qualifications or exclusions to the identification of the theme
5. Examples, both positive and negative, to eliminate possible confusion when looking for the theme (Boyatzis, 1998, pp. x–xi).

Thematic coding makes distinction between manifest (themes directly observable in the information) and latent coding (themes underlying the phenomenon). Manifest content analysis can be considered the content analysis of the visible or apparent content of a phenomenon. Latent content analysis, on the other hand, is looking for the underlying aspects of the phenomenon under investigation (Boyatzis, 1998, p. 16). Potter and Levine-Donnerstein (1999) note that for latent content the coders must provide subjective interpretations based on their own mental schema and that this “only increases the importance of making the case that the judgments of the coders are intersubjective, that is, those judgments, while subjectively derived, are shared across coders, and the meaning therefore is likely to reach out to readers of the research” (p. 226). According to the authors, with manifest content, the coding task is one of clerical recording and objectivity can be achieved relatively easily. Objectivity is much more difficult to achieve with latent content. Thematic analysis enables the researcher to use both manifest and latent content analysis at the same time.

Table 4-1 illustrates how the characteristics of a good code applied to the construct of authentic leadership (Table 4-2).

The outcome of the coding process, whether performed qualitatively or quantitatively is a codebook which includes the following elements: (1) a code mnemonic or label, a code definition, a full definition of inclusion criteria, a full definition of exclusion criteria to explain how the code differed from others, and example passages that illustrate how the code concept may appear in natural language. Though the structure is simple and appears to be stable, the process of building a codebook

Table 4-1 Characteristics of a good code using Boyatzis' criteria

-
1. *What am I going to call (label) it?* Authentic leadership
 2. *How am I going to define it?* Leaders who are true to themselves
 3. *How am I going to recognize it in the data?* When respondents explicitly say they are authentic leaders who are deeply aware of how they think and behave and are perceived by others as being aware of their own and others' values/moral perspectives, and of high moral character (Avolio & Gardner, 2005)
 4. *What do I want to exclude?* Authenticity stems from self-awareness, knowing one-self, one's strength, and weaknesses. Attributions of authentic leadership to external forces or socialization agents do not qualify as authentic leadership
 5. *What is an example?* "Boards should choose authentic leaders for character, not charisma, for their values, and for their ability to motivate employees to create genuine value for customer"
-

Source: *Klenke (2008)*.

Table 4-2 Thematic analysis definitions

Thematic analysis is a process for encoding qualitative information. The encoding requires an explicit "code."

A theme is a pattern found in the information that at a minimum describes and organizes the possible observations or at a maximum interprets aspects of the phenomenon. A theme may be identified at the manifest level (directly observable in the information) or at the latent level (categorizing issues underlying the phenomenon).

A code may be a list of themes; a complex model with themes, indicators, and qualifications that are causally related; or something in between these two forms.

A codebook is the compilation or integration of a number of codes in a study.

Three methods for developing themes are:

1. Initially generating themes inductively from raw information
 2. Initially generating themes deductively from theory
 3. Initially generating themes from prior research
-

Source: *Boyatzis (1998, p. 161)*.

is dynamic. Relational database management programs such as ACCESS® (Microsoft Corporation) are suited for maintaining the structure and supporting the process. In a relational database, the basic unit is the relation, which is generally conceptualized as a table with rows of codes and columns of attributes or the definitional parameters of the code. Unlike many database management systems that are specifically designed for quantitative data, the cells in a table in ACCESS can contain unlimited amounts of text.

From Codes to Categories

Codes are sorted into categories based on how different codes are related and linked (Patton, 2002). They can be derived inductively and deductively and again are driven by the research questions and the knowledge about the phenomenon that is being investigated that the content analysts bring to the study. They may also be created theoretically or empirically. In general, the more theory-oriented the research, the more extensive the researcher's previous knowledge, the more focused the research question, and the more specific the existing hypotheses, the easier it is to move from codes to categories (Kuckartz, 2014). The basic goal in the coding process is to organize a large body of text into much fewer content categories, that is, data reduction. As analysis proceeds, additional codes are developed, and the initial coding scheme is revised and refined. Categories may be defined based on the research question and/or underlying theories and constructs. For example, in the opening vignette, the authors (Sultana et al., 2013) established cultural diversity management, conducive work environment, and competitive advantage as categories each of which subsumed a number of codes or subcategories.

Data analytic procedures depend on whether the content analyst employs a quantitative or qualitative approach. At one end is an exclusively qualitative and verbally descriptive, narrative approach to the phenomenon under investigation; at the other end is a primarily quantitative approach of statistically analyzing the phenomenon under investigation. For quantitative approaches, data to be analyzed may be frequency counts and lists of categories for which a variety of statistical procedures including correlations, *t*-tests, and structural equation models are available. For example, Bligh et al. (2004b) in a content analysis of the rhetorical content of President George Bush's speeches before and after the terrorist attacks of 9/11 reported means, standard deviations, and intercorrelations between the categories the authors had identified and additionally conducted a one-way multivariate analysis of covariance (MANCOVA) to examine if there were significant differences in Bush's speech before and after 9/11 on the variables of interest.

On the other hand, if the content analyst pursues a purely qualitative analytic strategy usually does not standardize the corpus of text and is more likely to weave quotes from the analyzed texts into the analysis and interpretation. As noted earlier, Boyatzis (1998) developed a procedure known as thematic analysis which allows the development of codes in a purely inductive fashion. For example, in the analysis of life stories of leaders, the content analyst may mark key activities, competencies, obstacles to success, thoughts, and emotions as potential themes in the outline of

these life stories. However, descriptive or interpretive analyses do not preclude scoring or scaling of themes and if the analyst has a sufficient corpus of text, thematic analysis allows him or her to translate the qualitative codes into numeric themes or empirically creating clusters of themes.

Major sources of bias in content analysis are the researcher assumptions and the selection of texts to be analyzed. In addition, the development of categories is not a neutral or value free phase of the research process. Instead, as [Insch et al. \(1997\)](#) note, category development involves the similar researcher biases that are also present in the development of questionnaires. Additionally, the organization of the data researchers produce has considerable bearing on the interpretive process and serves as a source of bias as well. Bottlenecks refer to aspects of the analytic work that take the most time for the least analytical payoff. Data coding is the most time- and effort-consuming stage of content analysis because it involves considerable effort on part of the researcher, that is, getting to know the data, identifying recurrent concepts, abstracting constructs from the data, and building or applying theoretical models.

Use of Software in Content Analysis

Computer programs have been developed over the last 30 plus years to support content analysis and the interpretation of textual data and represent an alternative tool to the pen/paper/highlighter/scissors approach previously used ([Carcary, 2011](#)). Early on, [Fielding and Lee \(1998\)](#) point to the different functions software programs can serve in content analysis that include: (1) computer programs assist the content analyst by working through the material, underlining, writing marginal notes, defining category definitions and coding rules, and recording comments on the material. Most software programs offer helpful tools for handling text, searching, collecting, and editing passages; (2) computer programs function as documentation center, recording all steps of analysis of all interpreters thereby facilitating replication; (3) computer programs offer links to quantitative analyses by allowing data transfer to other programs such as SPSS. While initially serving primarily as search and retrieve a tool, the functionality of CAQDAS software has been greatly expanded. Most CAQDAS usually encompass some or all of the following tools: content searching tools, linking tools, coding tools, query tools, writing and annotation tools, and mapping or networking tools. Some such as NVivo can also be used in literature reviews, especially potentially identifying gaps in the literature and report writing ([Wickham & Woods, 2005](#)). Technological developments have significantly changed the ways we collect, transcribe, analyze, and interpret data ([Evers, 2011](#)).

Working with content analyses, among the software programs used are Atlas.ti, NNivo, NUD*ST and winMAX. Atlas.ti, for example, is a powerful software workbench for qualitative analysis of large bodies of textual, graphical, audio, and video data while winMAX is a tool for textual analysis and enables the combination of both qualitative and quantitative analytical procedures. These programs are particularly useful in the initial stages of the analysis since most of them provide a so-called “auto-code” functionality, but it is typically limited and only applicable to the level of words or basic patterns, not concepts. Qualitative analysis software also facilitates codebook development, the production of multiple versions of a codebook, the generation of reports that lists coding guidelines and changes the analyst made to them in the coding process and the preparation of outputs in a variety of formats including ASCII text files, spreadsheets, and other database formats.

Computerized content analysis has both strengths and limitations. [Morris \(1994\)](#), in the earlier days of CAQDAS, pointed out a number of advantages, including (a) perfect stability of the coding scheme, (b) explicit coding rules yielding comparable results, (c) perfect reliability (freeing the researcher to focus on issues of validity, interpretation, and explanation), (d) easy manipulation of the text to create output such as frequency counts and key-word in-contexts listings, and (e) the ability to easily uncover co-occurrences of important concept.

However, despite its strength, a number of limitations of computerized content analysis have been described as well. They include (a) a lack of natural language processing capabilities, (b) an insensitivity to linguistic nuances such as negation, irony, and tone, (c) the inability of the content analyst to provide a completely exhaustive listing of key words, (d) the inability of software to resolve references back and forth to words elsewhere in the text, and (e) the danger of word crunching or transforming rich meanings into meaningless numbers ([Morris, 1994](#)). The author concluded that computerized and human content analysis may be equally effective, but differences in results may occur at the higher levels of analysis (i.e., the paragraph or the whole document) in which humans are more sensitive to larger contextual cues.

Later discussions of the limitations of CAQDAS (e.g., [Davis & Meyer, 2009](#)) noted that even with the best content analysis software tools, the researcher is still faced with a considerable amount of manual labor. Current programs also do not relieve the content analysis from the hard work of generating categories and themes. In the final analysis, it is the researcher who conducts the analyses, not a computer program. In addition, humans are more prone to errors than machines, especially when fatigue sets in. Additionally, media content analysis including screen scraping of TV data, Internet focused text research, and graphical visualization of analytical

Table 4-3 Advantages and disadvantages of CAQDAS

Advantages	Disadvantages
1. Stability of coding scheme	1. Danger of word crunching, transforming rich meanings into meaningless numbers
2. Explicit coding rules yielding comparable results	2. Insensitivity to linguistic nuances such as negation, irony, and tone
3. Easy manipulation of text to create output such as frequency counts	3. Inability of researchers to provide a completely exhaustive listing of key words
4. Ability to easily uncover co-occurrences of important concepts	4. Inability of software to resolve references back and forth to words elsewhere in text
5. Easy support of multidisciplinary research teams	5. Software programs do not conduct analyses, researchers do
6. Simple coding	6. Lack of software-related flexibility

Source: *Adapted from Paulus, Lester, and Britt (2013), Davis and Meyer (2009), Flick (2009).*

relationships encoded in the text are poorly served by existing software. Finally, software applications are not neutral tools (Table 4-3).

Among the new developments in CAQDAS that have yet to find their ways into leadership research is the growing trend toward cloud computing, the proliferation of mobile devices based on portable computing power used by iPhones, tablets, wireless interviewing, and viral sampling through social media sites such as Facebook and Twitter (Palys & Atchison, 2012). Flick (2009) concluded that CAQDAS programs should not be “overloaded with hopes and expectancies, nor should they be demonized” and in general “no technology revolution of qualitative research” has yet occurred (p. 370). Thompson (2002) pointed out that in many qualitative studies involving CAQDAS software much is left unsaid as though the very name of the computer program in itself is sufficient to justify the way the data are analyzed.

QUALITY CRITERIA AND VALIDATION ISSUES

Intercoder Agreement in Content Analysis

Interrater reliability, and more specifically, intercoder agreement, is a critical component of content analysis because if the coding is not reliable, the analysis cannot be trusted (Krippendorff, 2012). It shows to what extent different coders agree in the coding of the same text. Because of the problems of reliability, the coding of texts is usually assigned to multiple coders so that the researcher can determine if

the constructs being investigated are shared and if multiple coders can reliably apply the same codes. As Neuendorf (2002) notes, “given that a goal of content analysis is to identify and record relatively objective (or at least intersubjective) characteristics of messages, reliability is paramount. Without the establishment of reliability, content analysis measures are useless” (p. 141). High levels of disagreement among coders suggest weakness in research methods, including the possibility of poor operational definitions, categories, and coder training. There are also important practical reasons to establish intercoder reliability. Neuendorf (2002) argues that, in addition to being a necessary (although not sufficient) step in validating a coding scheme, a high level of reliability also has the practical benefit of allowing the researcher to divide the coding work among different coders.

Measuring intercoder reliability

Intercoder reliability is assessed by having two or more coders categorize units of analysis (articles, stories, words, speeches, etc.) and then using the categorizations to calculate a numerical index agreement between or among the coders (Krippendorff, 2004a; Neuendorf, 2002; Patton, 2002). In classical content analysis, appropriate coefficients are calculated to determine intercoder agreement to include: (1) percentage agreement; (2) Scott’s π (π); (3) Cohen’s kappa (κ); and (4) Krippendorff’s alpha (α) (Krippendorff, 2004a, 2004b).

Percentage agreement is based on the percentage of all coding decisions made by a pair of coders on which the coders agree. Percentage agreement takes values ranging from .00 (no agreement) to 1.00 (perfect agreement). The obvious advantage of this statistic is that it is simple, intuitive, and easy to calculate. It can also accommodate any number of coder. The major weakness of percentage agreement is that this method fails to account for agreement that would simply occur by chance. For example, assume that two coders are given 100 news stories to code as making reference or not making reference to leadership success. Without any instructions or training, even without knowing the concept they are to code, they will agree half of the time, and these random agreements will produce a percentage agreement value of .50. Characteristics of percentage agreement also allow researchers to artificially inflate reliability by adding categories they know will rarely be used or produce disagreement. Another limitation is that percentage agreement only records agreements and disagreements and gives no credit for coders whose decisions are “close” (Cohen, 1968).

Scott’s π is intended to control for the role of chance agreement and uses a joint distribution across two coders. As Neuendorf (2002) points out, this takes into account not just the number of categories but how these categories are used by the coders. The normal range of this statistic is from .00 (agreement at chance level) to 1.00

(perfect agreement), and a value of less than .00 indicates agreement less than chance. Scott's π assumes nominal level data, is only appropriate for two coders and ignores differences in how the two coders distribute their evaluations across coding categories (Scott, 1955). In other words, this index does not account for the differences in how the individual coders distribute their values across the coding categories, a potential source of systematic bias. It assumes that coders have distributed their values across the categories identically and if this is not the case, the formula fails to account for the reduced agreement (Neuendorf, 2002).

Cohen's κ (1960, 1968) index also accounts for chance agreement using the same conceptual formula as Scott's π . This statistic was intended as an improvement over Scott's π , taking into account the differences in coders' distributions by using a multiplicative term instead of an additive one (Cohen, 1960). More specifically, expected agreement by chance in this case is calculated based on the "multiplicative marginals" rather than additive ones, which has the effect of accounting for differences in the distribution of values across the categories of different coders. Cohen (1968) proposed a weighted κ to account for different types of disagreements. Like π , it assumes nominal level data and has a normal range from .00 (agreement at chance level) to 1.00 (perfect agreement), and a value of less than .00 indicates agreement less than chance. Both π and κ have been criticized as being overly conservative, giving credit only to agreement beyond chance, a tough challenge in the case of extreme distributions (Potter & Levine-Donnerstein, 1999).

Krippendorff's α , originally formulated in 1980, takes into account chance agreement and, in addition, the magnitude of the misses, adjusting for whether a variable is measured as nominal, ordinal, interval, or ratio. The author (Krippendorff, 2004a) claims that this index is the most general agreement measure with appropriate reliability interpretations in content analysis. It can be applied to a variety of data and any number of coders, is applicable to small and large sample sizes alike, and is using the same assumptions as Scott's π of equal marginal proportions for the coders. The biggest drawback to its use has been its complexity and the resulting difficulty of performing hand calculations, especially for interval and ratio level variables. However, recently "macros" or customized programming that can be used with existing software to automate calculations have become available in statistical programs such as SPSS as well as stand-alone software such as ProGAMMA (2002).

A number of rules of thumb for standards of intercoder reliability have been proposed. For example, in many textbooks on qualitative research methods (e.g., Patton, 2002) 75–80 percent agreement or intercoder correlations of .70–.80 and higher are suggested as indicative of high reliability. Krippendorff (1980), for

example, without specifying the type of reliability, proposed a rule of thumb if intercoder reliability is above .80 is acceptable, with only “highly tentative and cautious conclusions” made about variables with reliabilities between .67 and .80. Neuendorf (2002) reviews rules of thumb set out by several methodologists and concludes that intercoder “coefficients of .90 or greater would be acceptable to all, .80 or greater would be acceptable in most situations, and below that, there exists great disagreement” (p. 145). The criterion of .70 is often used for exploratory research. Higher criteria should be used for indices known to be liberal (i.e., percent agreement) and lower criteria can be used for indices known to be more conservative (Cohen’s κ , Scott’s π , and Krippendorff’s α).

In contrast, qualitative procedures for establishing intercoder agreement are based on procedures that are aimed at resolving intercoder disagreements through discussions and conflict resolution where disparities in codes and categories exist.

Threats to reliability in content analysis may result from a number of conditions including: (1) a poorly executed coding scheme; (2) inadequate coder training; and (3) coder fatigue. The following recommendations (Krippendorff, 2004b; Neuendorf, 2002) to enhance reliability in content analysis have been offered: (1) reliability data defined as the sample of data from which the trustworthiness of a population of data is to be inferred, must be generated by coders who are widely available, follow explicit and communicable instructions (a data language), and work independently of each other; (2) a decisive agreement coefficient should measure agreements among multiple descriptions of each unit of analysis. For two coders, large sample sizes and nominal data, π is such a coefficient. Krippendorff’s α can handle multiple coders, all scales of measurement, missing data, and small sample sizes; (3) an acceptable level of agreement below which data are to be rejected as too unreliable must be chosen depending on the costs of drawing invalid conclusions from these data; (4) after data have been generated, reliability may be improved by removing unreliable distinctions from the data, recoding or lumping categories, or dropping variables that do not meet the required level of reliability; (5) a preferred approach is to calculate and report two (or more) indices, establishing a decision rule that takes into account the assumptions and/or weaknesses of each (e.g., to be considered reliable, a variable may be at or above a moderate level for a conservative index, or at or above a high level for a liberal index). In any case the researcher should be prepared to justify the criterion or criteria used; (6) do *not* use only percent agreement to calculate reliability. Despite its simplicity and widespread use, there is consensus in the methods literature that percentage agreement is a misleading and inappropriately liberal measure of intercoder agreement (at least for nominal-level variables); if it is reported at all, the researcher must justify its value in the context of the attributes of the data and analyses at hand.

Krippendorff (2004a, 2004b) discusses the following specific quality criteria for content analysis. *Semantic validity* relates to the meaning reconstruction of the material, and is expressed in the appropriateness of the category definitions, the key examples, and the rules for coders. Evidence on semantic validity ascertains the extent to which the categories of an analysis of texts correspond to the meanings these texts have within the chosen context. *Sampling validity* refers to the usual criteria for precise sampling. Ideally, content analysts actively sample a population, using sampling plans to ensure representativeness. But in many practical situations, texts become available by their sources' choice and contain unintentional biases in representing the phenomena of their interest. *Correlative validity* refers to the degree to which the findings obtained by one method correlate with findings obtained by other methods such as observation.

Predictive validity establishes the degree to which the answers of a content analysis accurately anticipate events, identify properties, or describe states of affairs. It can only be used as a quality criterion, if predictions can be reasonably be made from the material. *Construct validity* relates to previous success with similar constructs, established theories and models, and representative interpretations. Stability refers to whether the same results are obtained in a renewed application of the same analytical tool to the same text and reproducibility is the extent to which the analysis achieves the same results under different circumstances, for instance with different coders. It can be measured through intercoder reliability for which a range of indices have been developed. Finally, accuracy assumes stability and reproducibility and denotes the extent to which the analysis meets a particular standard.

SELECTED EXAMPLES OF CONTENT ANALYSIS STUDIES IN LEADERSHIP RESEARCH

As in the case of choosing a representative sample of case studies on leadership, a wide net was cast in selecting the content analysis studies depicted in Table 4-4. Included in sample of content analysis in leadership research are diverse topics including analysis of responses of corporate leaders to 9/11 attacks (Schwartz, 2013), analysis of leadership and ethics related language in scholarly versus popular publication, analysis of speeches of political leaders (e.g., Bligh et al., 2004b), leader emergence during the polish solidarity movement (Biezenski, 1996), barriers to women in leadership roles (Hall Lewis, 2012), studies conducted in the United States as well as abroad, and diverse methodologies to include both quantitative and qualitative content analysis techniques, manual and computerized coding. In this section, the goal is to highlight both the uniqueness and versatility of content analysis in leadership research. Because of page limitations, not all content analysis

Table 4-4 Selected examples of content analysis studies in leadership research

Author(s)	Purpose	Respondents	Data sources	Unit of analysis	Analyses	Reliability
Schwartz (2013)	To examine responses of corporate leaders to 9/11 attacks	Eight leaders of corporation hit in 9/11/2001	Books, newspaper articles, video	Phrases	Frequency of constructs	NR
Lewis (2012)	To analyze gender bias and underrepresentation of women leaders in higher education		Ten leadership books	Directive and summative content analysis	Qualitative; coding books using Bolman and Deals	
Saldivar et al. (2012)	To determine differences in use of leadership and ethics related language		432 scholarly articles, popular magazines, and newspaper articles	Words and phrases	Two analyses: (1) word and phrase counts; (2) multiple ANOVA followed by univariate tests	NR
Brocato and Theodori (2011)	To identify the operationalization of leadership traits and follower behaviors		17 peer reviewed journal articles published between 1998 and 2008	Phrases	Qualitative content summary of transformational and charismatic traits	NR

Table 4-4 (Continued)

Author(s)	Purpose	Respondents	Data sources	Unit of analysis	Analyses	Reliability
Bligh et al. (2004b)	To explore how elements of President Busch's speeches changes in response to 9/11		74 speeches and radio addresses	Speeches	Quantitative: descriptive statistics, intercorrelations, ANCOVA	Kendall's coefficient of concordance
Boulais (2002)	To examine the applicability of Kouzes and Posner's leadership challenge to selected children's books	NA	Selected children's books	NR	Qualitative	NR
Strange and Mumford (2002)	To test for a distinction in the behaviors of ideological (personalized) and charismatic (socialized) leaders	Visionary leaders	Biographies, historic outcome data	Word, phrase, sentence	Quantitative: multivariate, univariate analyses, discriminant analysis, correlation coefficients	Cohen's kappa
Kirkpatrick, Wofford, and Baum (2002)	To demonstrate the feasibility of the motive coding methodology applied to vision statement	Wood working firms; federal engineering services agency	Vision statements		Qualitative: thematic coding of texts	Descriptive statistics, correlations, regression
Brown and Barker (2001)	To identify components of the Hersey-Blanchard Situational Leadership Model framework	Managers of Fortune 500 financial services company	Responses to open-ended survey questions	Sentence phrase	Quantitative: percentages	Cohen's kappa

Buttner (2001)	To determine female entrepreneurs' roles in organizations using relational theory	129 female entrepreneurs	Focus group transcripts	Sentence phrase	Computer-assisted analysis	Test–retest reliability (87.9%) with 5 months interval between coding procedures
Golan and Wanta (2001)	To determine how newspapers framed Bush and McCain during the New Hampshire Republican primary	Journalists	Articles from regional newspapers; gallup reports	Paragraph	Quantitative: frequencies, chi-square tests, comparison to gallop data	Not reported
Bystrom, Robertson, and Banwart (2001)	To explore the media's portrayal of female candidates in comparison to male candidates in primary races for governor and U.S. Senate offices	Journalists	Articles from major national and regional newspapers	Article	Quantitative: descriptive statistics, frequencies, squares, paired-sample <i>t</i> -tests	Holsti intercoder reliability
Awamleh and Gardner (1999)	To examine the effects of vision content, delivery, and organizational performance on perceptions of leader charisma	Actors	Video-taped speeches	Rhetorical theme	Quantitative: frequencies, varimax rotation	NR

NR = not reported.

studies depicted in [Table 4-4](#) are reviewed because of some overlap in either content or procedures or in favor of more recent studies. For example, Brocato's and Theodori's content analysis of transformational and charismatic leadership concepts from a sample of journal articles is similar to the Hodgson et al. article described in more detail, except that the analysis were conducted qualitatively only.

[Schwartz \(2013\)](#) content analyzed books, newspaper articles, and videos documenting the responses of eight corporate leaders to the terrorist attacks of 9/11/2001 using [Goleman's \(1995\)](#) theory of emotional intelligence consisting of five constructs: self-awareness, self-regulation, self-motivation, empathy, or relationship management. The men were in leadership positions and housed in the World Trade Center (WTC) on the day of the attacks and were highly visible in the press after the tragedy. Of the six companies represented by the eight leaders, Cantor Fitzgerald suffered the highest number of casualties which resulted in extensive media coverage. CEO Howard Lutnick also made appearances on national news networks to discuss the plight of his company. As such, more information was available on Cantor Fitzgerald than for the other companies while others that were located below the impact zone lost which were reduced to 90 phrases only one employee, resulting in less extensive media coverage.

Initially, the author coded 100 phrases in the first coding cycle that were used in the final analysis. The data were analyzed using constructs frequencies as they appeared in the corpus of text which included 18 books. The results revealed that all five of Goleman's construct were coded and identified in the leaders' responses. Three constructs proved particularly prevalent, empathy (38 percent of the findings), self-regulation (31 percent), and relationship management (19 percent); self-awareness (7 percent) and self-motivation (5 percent) were less common. A response made by Howard Lutnick coded for Goleman's emotional construct of empathy stated:

And at the center of this were the families. We were supporting each other, sharing information. Maybe I was helping them, but certainly it was helping me to talk with the families. Every time I left the crisis center I was energized. And I would describe it at the time as physical. Hugging, holding, and crying with the families energized me. I didn't need to sleep ([Schwartz, 2013, p. 74](#)).

Despite suffering immense losses, some of the companies were able to successfully rebuild. In addition to the complete destruction of their WTC offices, the four companies lost a combined total of 792 employees. Leaders of these companies led their organizations to overcome seemingly insurmountable odds, and each company thrives over a decade later under the same leadership. For example, Cantor

Fitzgerald resumed operations on September 13, 2001, only two days after the attacks (Gordon, 2001). In 2004, the company moved to their current offices in Midtown Manhattan. Each year, on September 11, the company holds a Charity Day in which 100 percent of the day's revenues are donated to various causes, raising tens of millions of dollars in the past decade (Petrecca, DiBlasio, & Dorrel, 2012). In the companies that triumphed over tragedy demonstrating remarkable adaptability and expressing deep empathy, the WTC leaders helped their companies through one of the most demanding leadership challenges imaginable. The events of 9/11 drastically transformed both the climate of the country and the context of the surrounding presidency (Bligh et al., 2004b).

Leadership and ethics are two terms used frequently in leadership theory and practice. Using content analysis, Hodgson, Green, and Kodatt (2012) analyzed 203 scholarly, peer reviewed articles, 166 newspapers, and 63 magazine articles (for a total corpus of text of 432 pages) to determine if there are significant differences in the interpretation of leadership and ethics language between two types of articles—scholarly, peer-reviewed publications and articles that appeared in the popular press published between 2000 and 2011. In addition to the words leadership and ethics, the authors also included language related to these terms such as transformational leadership consideration and management. The researchers reviewed the use of words and phrases electronically to verify that the use of the meaning units (phrases) was congruent with the categories in which they were placed. This was the qualitative phase of the analysis. In the subsequent quantitative phase, they conducted a two way analysis of covariance to determine whether there were differences between the seven themes they had established for the two types of publications: scholarly versus popular articles that appeared in newspaper and magazines. The results indicated significant differences for each of the themes derived from qualitative coding. For example, the univariate Scheffé post hoc analysis for the ethics and morality language showed that there was a higher percentage of ethics and morality-oriented language in leadership ethics articles than in the popular magazine or newspaper leadership articles. The authors concluded that researchers, as opposed to reporters who write for the popular press view leadership ethics through different lenses.

Hall Lewis's (2012) content analysis highlights both a controversial topic and an analytical techniques not used in other content analysis studies. The researcher set out to determine the degree of gender bias and underrepresentation of women in higher education using 10 leadership books and Bolman and Deal's (2008) four frame model of organizations that describes four frames: structural, human resource, symbolic, and political. The author then coded the four frames as either masculine or feminine based on the literature regarding masculine and feminine

leadership styles identifying the political and structural frames as masculine and the human resource and symbolical frames as feminine.

The method of analysis used in this research was a combination of directed and summative analysis, a qualitative approach. In the directed phase, the researcher developed a coding scheme based on current theory (Hsieh & Shannon, 2005), in this case Bolman and Deal's theory (2008) of leadership and organizations. The researcher codes the data based on the initial coding schemes and develops conclusion to support or refute the initial theory. In summative content analysis, the researcher utilized word count to determine how frequently certain words are used within a text. The researcher then examined the text for how words are used and derived meaning from the context (Hsieh & Shannon, 2005). Gender biases in texts can be determined in many different ways such as by pronoun choice, sentence structure, qualifying words, differentiation, examples, and slang to determine to what extent gender bias exists within the selected texts.

In the 10 leadership books included in this study, the author reported 3097 references to gender in the six ways that I measured gender bias—pronoun, example, differentiation, sentence structure, qualifying words, and slang. The overwhelming majority of instances were male gender references—2167 instances or 69.97 percent. Female gender references account for just 621 instances or 20.05 percent and neutral references such as the use of she/he accounted for 309 instances or 9.98 percent. For every female gender reference included in the leadership texts, the authors of the texts used 3.5 male gender references. The use of pronouns and examples of prominent leaders were the largest categories indicated of gender bias. The total number of examples, for instance, found within the 10 included texts was 2009. Of that total, 1596 (79.44 percent) were male examples, 389 (19.36 percent) were female examples, and 24 (1.19 percent) were neutral examples. Male examples outnumber female examples of 4:1. In Collins (2001) *Good to Great* highly popular book, male examples are used almost exclusively. Lewis found 236 male examples and just 16 female examples in this text. This book describes how 11 companies were able to achieve extraordinary success, as judged on the basis of the author's metrics, which included stock market performance and sustained profitability. The leaders of all 11 companies at the time of their inclusion in the book were men. In sum, women were noticeably underrepresented throughout the texts, as evidenced by the number of instances of male gender bias. However, feminine leadership styles were more prominent than masculine leadership styles. Therefore, while women and men were not included equally within the texts, feminine voices and perspectives were present which may represent a shift in culture.

Two international organizations, the Asia-Pacific Economic Corporation (APEC) and the Pacific Economic Corporation Council (PEEC) provided the context for the [Chiu and Lao \(2008\)](#) content analysis of leadership practices and organizational traits through an examination of major documents of these two organizations: the APEC Leaders' Declaration and PECC's reports from the bi-annual meetings. This content analysis was motivated by the numerous criticism leveraged against the two organizations including their ability to respond effectively to regional crises ([Lee, 2006](#)) and inefficient operational management ([Low, 2006](#)).

The authors used both manifest coding and latent coding. In manifest coding, the researchers identified particular verbs related to different leadership styles while latent coding involved interpreting the leadership implications based on the results of manifest coding. Examples of verbs that were identified in the manifest coding phase were commit, instruct, encourage, and support which the authors considered as indications of a high level of leadership commitment similar to directive leadership. In contrast, verbs such as recognize, acknowledge, or note were viewed as representing a low level of leadership commitment resembling [Bass \(1985\)](#) laissez-faire leadership style. The next step in the manifest coding was to determine the number of verbs relevant to the different levels of commitment appearing in each of the two documents. The latent coding method was then used to explain and compare the implied connotations of this content ([Chiu & Lao, 2008](#), p. 39). After identifying the leadership styles of the two companies, the authors then examined three organizational paradigms—rational, natural, and open systems—using a narrative approach.

The results revealed that among APEC leaders, directive leadership was the preferred leadership practice as seen by the use of 341 or 24.6 percent verbs from 1993 to 2006 compared to 228 verbs or 16.20 percent per declaration. APEC leaders used laissez-faire leadership 106 times per declaration. Thus, directive leadership was the preferred style in this organization. In the PECC organization, during the same time period, the company leaders used transformational leadership verbs 22 times or an average of 2.44 percent at each of the meetings. Laissez-faire leadership verbs were used 1.22 percent. Overall, the results of preferred leadership practices were the same for APEC and PECC indicative of a combination of directive, transformational, and laissez-faire leadership styles. However, in APEC directive leadership was practiced more often than in PECC more. The results from the organizational paradigm analysis are not reported here.

Analysis of speeches of political leaders is a major topical stream in content analysis research. Presidential political leadership was examined in presidential speeches of George Bush, Jesse Jackson, Bill Clinton, and Robert Dole (e.g.,

Bligh, Kohles, & Meindl, 2004a; Bligh et al., 2004b). Bligh et al. (2004a) examined the rhetoric of charismatic leadership as evidenced in Bush's speeches before and after 9/11 to determine if the context of crisis affected the President's communications with the American people (p. 211). The authors devised three studies (1) to determine whether the crisis of 9/11 caused a charismatic shift, that is, a move toward a more charismatic rhetorical style; (2) to examine the extent to which the crisis affected the media portrayal of and response to the President's rhetoric; and (3) to explore the relationship among the President's charismatic rhetoric, the media's portrayal of the President's speeches, and the public approval ratings over time. The authors noted that together these three studies afforded an opportunity to understand the shifting of mutual expectations regarding leadership that were precipitated by the events of 9/11.

For the three studies discussed in these articles, the authors used DICTION 5.0, a computerized content analysis specifically designed for political discourse. The program generated eight composite variables from the 33 dictionaries included in DICTION. Among these eight composite variables were a collective focus observed in speeches of charismatic leaders, confidence in followers' worth, similar to followers' and values and moral justification. Means, standard deviations, and intercorrelations were computed for each of the eight composite variables followed by MANCOVA to determine whether there were significant differences in Bush's speeches and whether this effect was also reflected in the media's response to the President's charismatic rhetoric before and after 9/11 on the variables of interest. In the third study, the sample comprised articles and transcripts from a variety of media outlets, television news, and programs. The authors then followed a similar procedure as in the first study utilizing DICTION to analyze each article.

Taken together, the results confirmed that the occurrence of the 9/11 crisis led to a change in Bush's rhetoric resulting in language that has been theoretically linked to charismatic leadership. Moreover, and in parallel, the media's portrayal of the President after 9/11 also incorporated more charismatic language (Bligh et al., 2004a, p. 229).

Boulais (2002) examined leadership in children's literature based on the Kouzes and Posner (1995) leadership framework by analyzing various works of children's literature for leadership themes. Kouzes and Posner (1995) developed a framework which became very popular designed both to understand leadership and as a tool for the development of leaders. The model rests on five leadership practices: The first leadership practice is called Challenging the Process, which, according to Kouzes and

Posner (1995), focuses on seeking out and accepting challenges as well as leaders who challenge the process by recognizing and supporting innovative ideas, and taking risks to bring about change. A children's book containing strong images and metaphors and examples of Challenging the Process is *A Story, A Story* (Haley, 1970), a story based on African folklore. It retells the story of Ananse the Spider man who challenges the process by first taking the risk of approaching the Sky God and asking for the golden box of stories. The second leadership practice, known as Inspiring a Shared Vision, refers to a leadership practice based on the leader's ability to imagine the future and clearly articulate the vision to others. A children's book rich in imagery reflecting this leadership practice is Ringgold's (1991) *Tar Beach*. The heroine of the book, a young girl name Cassie who grew up as a minority in Harlem. Cassie spends many summer nights on the roof of her apartment building from where she sets out on a visionary journey flying over the city. Ringgold notes that the imagery of flying draws on traditional African American folktales where flying is a metaphor of escaping slavery. In the end, the author portrays the importance of sharing a vision by showing how Cassie enlists her brother in the common vision of hope for a better life (Boulais, 2002, p. 58).

Enabling Others to Act, the third category in Kouzes and Posner's (1995) leadership framework is a shorthand description that includes leadership practices such as teamwork, building trust, and empowerment. A children's book containing strong imagery, metaphors and examples of this leadership practice is Ransome's (1968) retelling of *The Fool of the World and the Flying Ship*. This book is based on a Russian folktale in which a young man builds a flying ship and sails of to win the hand of the Czar's daughter. On the journey, the young man encounters and befriends a variety of unusual individuals like the Swift-goer and Eater who enable him to meet the challenges, a series of seemingly impossible tasks, the Czar had demanded to be completed before allowing him to marry the Princess. Modeling the Way, as defined by Kouzes and Posner (1995), describes the importance of leading by example. Finally, Encouraging the Heart captures a leadership practice based on small and large actions by a leader who encourages and supports the followers and helps the team to celebrate victories.

The results of Boulais's (2002) content analysis identified several children's book in this category. Musgrove's (1976) book *Ashanti to Zulu: African Traditions* illustrated by Leo and Diane Dillon is an alphabet book which draws on the alphabet theme to describe 26 distinct aspect of African culture such as ceremonies, celebrations and other customs. The authors are able to effectively utilize the format of this alphabet picture book to recognize and celebrate the individual contributions of multiple African tribes and cultures that together weave the tapestry of Africa (Boulais, 2002, p. 60).

Finally, Boulais, through content analysis identified several children's book that exemplify each of the five leadership practices outlined by Kouzes and Posner (1995). Wiesner's (1999) *Sector 7* is one example. In this story, a young boy on a class trip to the Empire State Building meets and is befriended by a small cloud. They travel together to Sector 7, which is the factory where all the clouds are created. In the story the clouds have a vision of changing the old cloud designs into fun shapes. When the young boy arrives at Sector 7, the clouds share their vision with the young boy. Inspired to help the clouds achieve their goals, the young boy enables the clouds to act by creating blueprints from their ideas. With these new designs in hand, the clouds challenge the process and tradition of old cloud making by becoming the new and innovative shapes. Despite the boy's expulsion from Sector 7, the idea of innovation, uniqueness, and thinking outside the box is modeled for the adult cloud designers. Likewise, the boys' support of the clouds' desire for uniqueness encourages their heart and celebrates their self-expression.

Although from a methodological perspective, this study is lacking rigor since Boulais (2002) does not describe sampling procedures that led to the identification of the children's books that were content analyzed nor were the specific content analysis steps and procedures described in the article, it is an innovative and creative study that shows a unique application of content analysis by using the well-established leadership model developed by Kouzes and Posner (1995) and analyzing the five model components in children's literature. In addition, many of the books included in the sample were illustrated allowing not only for textual analysis but also exploring the five leadership practices in imagery found in the illustration.

Buttner (2001) conducted a content analysis of female entrepreneur's leadership style using relational theory as the guiding conceptual framework. Numerous studies on female–male differences in leadership style (e.g., Eagly & Carli, 2007) reported that women leaders prefer a style that focuses on the quality of leader–follower interactions, emphasize information sharing, are more participative in decision-making and other “feminine” attributes that characterize women's leadership styles. The author described her study as exploratory, designed to capture the women's voices. The women entrepreneurs participated in focus group interviews in which they discussed their experiences as business owner and the way they enacted their role as leader in their organization. The interviews were videotaped and transcribed. An ethnographic software program, Ethnograph, was used for the content analysis of the focus group comments. All comments in the transcripts in which the entrepreneurs talk about relationships with or management of subordinates were assembled into a master file and then were sorted into four categories: (1) persevering characterized by a focus on task through nurturing, protecting, and

safeguarding; (2) mutual empowering characterized by a focus on developing another person such as a subordinate or client; (3) achieving characterized by using relational skills to enhance the entrepreneur's own professional growth and effectiveness; and (4) creating team characterized by a focus on creating a sense of team. Each of these four categories had a number of subdimensions (Buttner, 2001, p. 259). The results indicated that relational theory is a useful framework for examining the ways women entrepreneurs approach relationships in their business. Their management style was based on their values and beliefs about effective ways to relate to employees and clients outside a previously male dominated culture or tradition (p. 263).

The review of the content analysis cited in this section leads to three major observations. Charismatic and political leadership were the topic of numerous studies using content analyses (e.g., Bligh et al., 2004a, 2004b; Tan & Wee, 2002). Often followers are mesmerized and captivated by charismatic leaders' rhetoric such as Martin Luther King's "I have a dream" and J.F. Kennedy's "Ask not what your country can do for you. Ask what you can do for your country." Charismatic leaders tend to be spellbinding orators. Although leaders can be charismatic without being spellbinding orators, the ability to capture an audience through oratory plays an important part in the formation of charisma.

The second observation concerns the preponderance of quantitative (and combined quantitative/qualitative) content analyses which support the approach of proposed by Krippendorff (2014) and Neuendorf (2002). It also reflects the lingering hegemony of positivism in the study of leadership. Quantitative content analyses are more likely published in top tier journals such as the *Journal of Applied Psychology* or *The Leadership Quarterly*. The third observation involves the use of software programs to facilitate the task of word crunching and the results of coding can easily imported into statistical packages such as SPSS or SAS to perform univariate or multivariate analyses.

SUMMARY

This chapter introduced content analysis as a flexible and versatile qualitative research method frequently used in the study of leadership that can be applied to a wide range of topics such as charismatic and political leader or the underrepresentation of women in higher education. Textual data may be derived from a variety of sources including books, articles, newspaper headlines, historical documents, transcripts of interviews, and advertising but also includes nontext based materials such as TV segments, photography, video tapes, and materials captured from the

performing and visual arts or any other communicative medium. I described the fundamental procedures for conducting a content analysis, the cornerstones of which are coding procedures which to include qualitative, quantitative, and computerized methods. Content analysis does not proceed in a linear fashion but represents a family of data collection and analytic procedures, each with its distinctive characteristics that can be mapped from a purely qualitative to a purely quantitative approach. For quantitative content analysis, CAQDAS software programs greatly facilitate the word crunching activities yet leave the final interpretive burden on the human content analyst. Therefore, CAQDAS programs will not replace traditional content analysis methods, but they add a useful complementary and alternative method to confirm predicted themes or to identify unexpected ones. Ultimately it is the researcher who analyzes and interprets that data, not a computer program.