Three Approaches to Qualitative Content Analysis

Hsiu-Fang Hsieh
Sarah E. Shannon

Content analysis is a widely used qualitative research technique. Rather than being a single method, current applications of content analysis show three distinct approaches: conventional, directed, or summative. All three approaches are used to interpret meaning from the content of text data and, hence, adhere to the naturalistic paradigm. The major differences among the approaches are coding schemes, origins of codes, and threats to trustworthiness. In conventional content analysis, coding categories are derived directly from the text data. With a directed approach, analysis starts with a theory or relevant research findings as guidance for initial codes. A summative content analysis involves counting and comparisons, usually of keywords or content, followed by the interpretation of the underlying context. The authors delineate analytic procedures specific to each approach and techniques addressing trustworthiness with hypothetical examples drawn from the area of end-of-life care.

Keywords: *content analysis; qualitative research; research methodology; end-of-life care*

Content analysis is a research method that has come into wide use in health studies in recent years. A search of *content analysis* as a subject heading term in the Cumulative Index to Nursing and Allied Health Literature produced more than 4,000 articles published between 1991 and 2002. The number of studies reporting the use of content analysis grew from only 97 in 1991 to 332 in 1997 and 601 in 2002.

Researchers regard content analysis as a flexible method for analyzing text data (Cavanagh, 1997). Content analysis describes a family of analytic approaches ranging from impressionistic, intuitive, interpretive analyses to systematic, strict textual analyses (Rosengren, 1981). The specific type of content analysis approach chosen by a researcher varies with the theoretical and substantive interests of the researcher and the problem being studied (Weber, 1990). Although this flexibility has made content analysis useful for a variety of researchers, the lack of a firm definition and procedures has potentially limited the application of content analysis (Tesch, 1990).

The differentiation of content analysis is usually limited to classifying it as primarily a qualitative versus quantitative research method. A more thorough analysis of the ways in which qualitative content analysis can be used would potentially illuminate key issues for researchers to consider in the design of studies purporting to

AUTHORS' NOTE: We wish to express our gratitude to Drs. Pamela L. Jordan, Carol J. Leppa, and J. Randall Curtis for their feedback and support in writing this article.

QUALITATIVE HEALTH RESEARCH, Vol. 15 No. 9, November 2005 1277-1288
DOI: 10.1177/1049732305276687

use content analysis and the analytic procedures employed in such studies, thus avoiding a muddling of methods (Morse, 1991).

Our purpose in this article is to present the breadth of approaches categorized as qualitative content analysis. We have identified three distinct approaches: conventional, directed, and summative. All three approaches are used to interpret text data from a predominately naturalistic paradigm. We begin with a brief review of the history and definitions of content analysis. We then illustrate the three different approaches to qualitative content analysis with hypothetical studies to explicate the issues of study design and analytical procedures for each approach.

BACKGROUND ON THE DEVELOPMENT OF CONTENT ANALYSIS

Content analysis has a long history in research, dating back to the 18th century in Scandinavia (Rosengren, 1981). In the United States, content analysis was first used as an analytic technique at the beginning of the 20th century (Barcus, 1959). Initially, researchers used content analysis as either a qualitative or quantitative method in their studies (Berelson, 1952). Later, content analysis was used primarily as a quantitative research method, with text data coded into explicit categories and then described using statistics. This approach is sometimes referred to as quantitative analysis of qualitative data (Morgan, 1993) and is not our primary focus in this article. More recently, the potential of content analysis as a method of qualitative analysis for health researchers has been recognized, leading to its increased application and popularity (Nandy & Sarvela, 1997).

Qualitative content analysis is one of numerous research methods used to analyze text data. Other methods include ethnography, grounded theory, phenomenology, and historical research. Research using qualitative content analysis focuses on the characteristics of language as communication with attention to the content or contextual meaning of the text (Budd, Thorp, & Donohew, 1967; Lindkvist, 1981; McTavish & Pirro, 1990; Tesch, 1990). Text data might be in verbal, print, or electronic form and might have been obtained from narrative responses, open-ended survey questions, interviews, focus groups, observations, or print media such as articles, books, or manuals (Kondracki & Wellman, 2002). Qualitative content analysis goes beyond merely counting words to examining language intensely for the purpose of classifying large amounts of text into an efficient number of categories that represent similar meanings (Weber, 1990). These categories can represent either explicit communication or inferred communication. The goal of content analysis is "to provide knowledge and understanding of the phenomenon under study" (Downe-Wamboldt, 1992, p. 314). In this article, qualitative content analysis is defined as a research method for the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns.

To illustrate the possible applications of content analysis, we constructed hypothetical studies drawn from the area of end-of-life (EOL) research. Content analysis has been a popular analytic method in studies related to EOL care, an area of increasing emphasis as demonstrated by its inclusion as one of the five research themes supported by the National Institutes of Health, National Institute of Nurs-

ing Research (NINR) for 2003 ("Enhancing the End-of-Life Experience for Patients and Their Families," NINR, 2003).

CONVENTIONAL CONTENT ANALYSIS

Researcher X used a conventional approach to content analysis in her study (Table 1). Conventional content analysis is generally used with a study design whose aim is to describe a phenomenon, in this case the emotional reactions of hospice patients. This type of design is usually appropriate when existing theory or research literature on a phenomenon is limited. Researchers avoid using preconceived categories (Kondracki & Wellman, 2002), instead allowing the categories and names for categories to flow from the data. Researchers immerse themselves in the data to allow new insights to emerge (Kondracki & Wellman, 2002), also described as inductive category development (Mayring, 2000). Many qualitative methods share this initial approach to study design and analysis.

If data are collected primarily through interviews, open-ended questions will be used. Probes also tend to be open-ended or specific to the participant's comments rather than to a preexisting theory, such as "Can you tell me more about that?" Data analysis starts with reading all data repeatedly to achieve immersion and obtain a sense of the whole (Tesch, 1990) as one would read a novel. Then, data are read word by word to derive codes (Miles & Huberman, 1994; Morgan, 1993; Morse & Field, 1995) by first highlighting the exact words from the text that appear to capture key thoughts or concepts. Next, the researcher approaches the text by making notes of his or her first impressions, thoughts, and initial analysis. As this process continues, labels for codes emerge that are reflective of more than one key thought. These often come directly from the text and are then become the initial coding scheme. Codes then are sorted into categories based on how different codes are related and linked. These emergent categories are used to organize and group codes into meaningful clusters (Coffey & Atkinson, 1996; Patton, 2002). Ideally, the numbers of clusters are between 10 and 15 to keep clusters broad enough to sort a large number of codes (Morse & Field, 1995).

Depending on the relationships between subcategories, researchers can combine or organize this larger number of subcategories into a smaller number of categories. A tree diagram can be developed to help in organizing these categories into a hierarchical structure (Morse & Field, 1995). Next, definitions for each category, subcategory, and code are developed. To prepare for reporting the findings, exemplars for each code and category are identified from the data. Depending on the purpose of the study, researchers might decide to identify the relationship between categories and subcategories further based on their concurrence, antecedents, or consequences (Morse & Field, 1995).

With a conventional approach to content analysis, relevant theories or other research findings are addressed in the discussion section of the study. In Researcher X's study, she might compare and contrast her findings to Kübler-Ross's (1969) theory. The discussion would include a summary of how the findings from her study contribute to knowledge in the area of interest and suggestions for practice, teaching, and future research.

The advantage of the conventional approach to content analysis is gaining direct information from study participants without imposing preconceived catego-

TABLE 1: Hypothetical Research Study Using a Conventional Approach to Content Analysis—Researcher X's Study

Little is known about the emotional reactions of terminally ill patients who are receiving hospice care, possibly because of their reluctance to discuss death issues (Wilson & Fletcher, 2002). Some patients might feel relieved to have active therapy end, whereas others might feel afraid or even abandoned. Researcher X wanted to learn more about the emotional experiences of hospice patients to be able to address their needs more effectively. Because there was no existing theory to serve as a framework for her study, her research question was "What are the emotional reactions of terminally ill patients who are receiving hospice care?"

Based on her clinical experience, Researcher X suspected that the emotional reactions of patients who were new to hospice care differed from those who had been in hospice care for a longer period. She also suspected that those receiving home hospice care had different experiences from those receiving in-patient hospice care. Researcher X therefore decided to use a stratified sampling technique to ensure heterogeneity of the sample. The target sample size was 10 home hospice patients and 10 inpatient hospice patients, with 5 from each group being recruited within 48 hours of enrollment into hospice and 5 recruited 7 to 10 days following enrollment. In addition, the sample would include both men and women and both older and middle-aged people.

Prior to recruitment and data collection, the research procedures were approved for use with human subjects. Informed consent was obtained from all participants. Researcher X collected data through individual interviews using open-ended questions such as "What has it been like to be in hospice care?" followed by specific probes. All interviews were audiotape-recorded and transcribed verbatim.

Researcher X used content analysis to analyze the data. She began by reading each transcript from beginning to end, as one would read a novel. Then, she read each transcript carefully, highlighting text that appeared to describe an emotional reaction and writing in the margin of the text a keyword or phrase that seemed to capture the emotional reaction, using the participant's words. As she worked through the transcript, she attempted to limit these developing codes as much as possible. After open coding of three to four transcripts, Researcher X decided on preliminary codes. She then coded the remaining transcripts (and recoded the original ones) using these codes and adding new codes when she encountered data that did not fit into an existing code.

Once all transcripts had been coded, Researcher X examined all data within a particular code. Some codes were combined during this process, whereas others were split into subcategories. Finally, she examined the final codes to organize them into a hierarchical structure if possible.

In the findings, the emotional responses of hospice patients were described using the identified codes and hierarchical structure. In discussion of the findings, the results from this content analysis were compared and contrasted with Kübler-Ross's (1969) model to highlight similarities and differences.

ries or theoretical perspectives. Researcher X's study depicts a research question appropriate for this approach. Knowledge generated from her content analysis is based on participants' unique perspectives and grounded in the actual data. Her sampling technique was designed to maximize diversity of emotional reactions, and the analysis techniques were structured to capture that complexity.

One challenge of this type of analysis is failing to develop a complete understanding of the context, thus failing to identify key categories. This can result in findings that do not accurately represent the data. Lincoln and Guba (1985) described this as credibility within the naturalistic paradigm of trustworthiness or internal validity within a paradigm of reliability and validity. Credibility can be established through activities such as peer debriefing, prolonged engagement, persistent observation, triangulation, negative case analysis, referential adequacy, and member checks (Lincoln & Guba, 1985; Manning, 1997).

Another challenge of the conventional approach to content analysis is that it can easily be confused with other qualitative methods such as grounded theory

method (GTM) or phenomenology. These methods share a similar initial analytical approach but go beyond content analysis to develop theory or a nuanced understanding of the lived experience. The conventional approach to content analysis is limited in both theory development and description of the lived experience, because both sampling and analysis procedures make the theoretical relationship between concepts difficult to infer from findings. At most, the result of a conventional content analysis is concept development or model building (Lindkvist, 1981). For example, Researcher X might find that patients who are new to hospice care express worry about how their social obligations will be met (such as finding care for a pet), whereas patients who have been in hospice for long periods might express more anticipatory grief. Researcher X might compare her findings to those of Kübler-Ross (1969) and conclude that an additional emotional reaction to entering hospice care is the process of "tying up loose ends," which she might define as making both financial and social arrangements.

DIRECTED CONTENT ANALYSIS

Sometimes, existing theory or prior research exists about a phenomenon that is incomplete or would benefit from further description. The qualitative researcher might choose to use a directed approach to content analysis, as Researcher Y did (Table 2). Potter and Levine-Donnerstein (1999) might categorize this as a deductive use of theory based on their distinctions on the role of theory. However the key tenets of the naturalistic paradigm form the foundation of Researcher Y's general approach to the study design and analysis. The goal of a directed approach to content analysis is to validate or extend conceptually a theoretical framework or theory. Existing theory or research can help focus the research question. It can provide predictions about the variables of interest or about the relationships among variables, thus helping to determine the initial coding scheme or relationships between codes. This has been referred to as deductive category application (Mayring, 2000).

Content analysis using a directed approach is guided by a more structured process than in a conventional approach (Hickey & Kipping, 1996). Using existing theory or prior research, researchers begin by identifying key concepts or variables as initial coding categories (Potter & Levine-Donnerstein, 1999). Next, operational definitions for each category are determined using the theory. In Researcher Y's study, Kübler-Ross's (1969) five stages of grief served as an initial framework to identify emotional stages of terminally ill patients.

If data are collected primarily through interviews, an open-ended question might be used, followed by targeted questions about the predetermined categories. After an open-ended question, Researcher Y used probes specifically to explore participants' experiences of denial, anger, bargaining, depression, and acceptance. Coding can begin with one of two strategies, depending on the research question. If the goal of the research is to identify and categorize all instances of a particular phenomenon, such as emotional reactions, then it might be helpful to read the transcript and highlight all text that on first impression appears to represent an emotional reaction. The next step in analysis would be to code all highlighted passages using the predetermined codes. Any text that could not be categorized with the initial coding scheme would be given a new code.

TABLE 2: Hypothetical Research Study Using a Directed Approach to Content Analysis—Researcher Y's Study

Despite their wide acceptance and popularity, Kübler-Ross's (1969) five stages of grief (denial, anger, bargaining, depression, and acceptance) have not been confirmed through other research. In taking care of terminally ill patients, Researcher Y wondered how well Kübler-Ross's theory described his patients' experiences with imminent death. His research question was "How well does Kübler-Ross's model describe the emotional passages or journeys of patients who have been diagnosed with a terminal illness?"

Researcher Y designed a sampling plan to maximize the chance of recruiting participants at different stages. All participants were diagnosed with a terminal illness, but one third were recruited while receiving "last chance" forms of curative therapy, one third after they refused further curative therapy but were not enrolled in hospice care, and one third who were contemplating (or had recently made) the decision to enter hospice care. In addition, the sample was recruited for gender balance and diagnostic diversity, specifically both oncology and non-oncology diagnoses. The target sample size was 18 to 21 participants. Interviews were conducted with individuals using open-ended questions, such as "What has your emotional journey been since being diagnosed with this illness?" Specific probes were developed based on Kübler-Ross's model, such as Have you felt angry since your diagnosis? After institutional review board approval, informed consent from all participants was obtained. All interviews were audiotape-recorded and transcribed verbatim.

Researcher Y developed operational definitions of the five emotional responses (anger, bargaining, etc.) identified in Kübler-Ross's model. He then reviewed all transcripts carefully, highlighting all text that appeared to describe an emotional response. All highlighted text was coded using the predetermined categories wherever possible. Text that could not be coded into one of these categories was coded with another label that captured the essence of the emotion. After coding, Researcher Y examined the data for each category to determine whether subcategories were needed for a category (e.g., anger toward self, anger toward doctors, anger toward spiritual being). Data that could not be coded into one of the five categories derived from the theory were reexamined to describe different emotional reactions. Finally, Researcher Y compared the extent to which the data were supportive of Kübler-Ross's theory versus how much represented different emotional responses. The report of study findings described the incidence of codes representing the emotional stages suggested by Kübler-Ross with those that represented different emotional responses by comparing the rank order of all codes. In the discussion section, Researcher Y summarized how the study validated Kübler-Ross's model and what new perspectives were added.

The second strategy that can be used in directed content analysis is to begin coding immediately with the predetermined codes. Data that cannot be coded are identified and analyzed later to determine if they represent a new category or a subcategory of an existing code. The choice of which of these approaches to use depends on the data and the researcher's goals. If the researcher wants to be sure to capture all possible occurrences of a phenomenon, such as an emotional reaction, highlighting identified text without coding might increase trustworthiness. If the researcher feels confident that initial coding will not bias the identification of relevant text, then coding can begin immediately. Depending on the type and breadth of a category, researchers might need to identify subcategories with subsequent analysis. For example, Researcher Y might decide to separate anger into subcategories depending on whom the anger was directed toward.

The findings from a directed content analysis offer supporting and nonsupporting evidence for a theory. This evidence can be presented by showing codes with exemplars and by offering descriptive evidence. Because the study design and analysis are unlikely to result in coded data that can be compared meaningfully using statistical tests of difference, the use of rank order comparisons of frequency

of codes can be used (Curtis et al., 2001). Researcher Y might choose to describe his study findings by reporting the incidence of codes that represented the five main categories derived from Kübler-Ross (1969) and the incidence of newly identified emotional reactions. He also could descriptively report the percent of supporting versus nonsupporting codes for each participant and for the total sample.

The theory or prior research used will guide the discussion of findings. Newly identified categories either offer a contradictory view of the phenomenon or might further refine, extend, and enrich the theory. In Researcher Y's study, the discussion might focus on the extent to which participants' emotional journeys paralleled Kübler-Ross's (1969) model and the newly identified emotional reactions or stages that were experienced by participants in the study.

The main strength of a directed approach to content analysis is that existing theory can be supported and extended. In addition, as research in an area grows, a directed approach makes explicit the reality that researchers are unlikely to be working from the naive perspective that is often viewed as the hallmark of naturalistic designs.

The directed approach does present challenges to the naturalistic paradigm. Using theory has some inherent limitations in that researchers approach the data with an informed but, nonetheless, strong bias. Hence, researchers might be more likely to find evidence that is supportive rather than nonsupportive of a theory. Second, in answering the probe questions, some participants might get cues to answer in a certain way or agree with the questions to please researchers. In Researcher Y's study, some patients might agree with the suggested emotional stages even though they did not experience the emotion. Third, an overemphasis on the theory can blind researchers to contextual aspects of the phenomenon. In Researcher Y's study, the emphasis on Kübler-Ross's (1969) stages of emotional response to loss might have clouded his ability to recognize contextual features that influence emotions. For example, the cross-sectional design of the study might have overemphasized current emotional reactions. These limitations are related to neutrality or confirmability of trustworthiness as the parallel concept to objectivity (Lincoln & Guba, 1985). To achieve neutral or unbiased results, an audit trail and audit process can be used. In Researcher Y's study, the vague terminology used in Kübler-Ross's description of the model would be a challenge for the researcher in creating useful operational definitions. Having an auditor review and examine these definitions before the study could greatly increase the accuracy of predetermined categories.

SUMMATIVE CONTENT ANALYSIS

Typically, a study using a summative approach to qualitative content analysis starts with identifying and quantifying certain words or content in text with the purpose of understanding the contextual use of the words or content (Table 3). This quantification is an attempt not to infer meaning but, rather, to explore usage. Analyzing for the appearance of a particular word or content in textual material is referred to as manifest content analysis (Potter & Levine-Donnerstein, 1999). If the analysis stopped at this point, the analysis would be quantitative, focusing on counting the frequency of specific words or content (Kondracki & Wellman, 2002). A summative approach to qualitative content analysis goes beyond mere word counts to include latent content analysis. Latent content analysis refers to the process of interpretation

TABLE 3: Hypothetical Research Study Using a Summative Approach to Content Analysis—Researcher Z's Study

Talking about death has virtually been banished from our language (Callahan, 1995). Use of the terms *die*, *dying*, and *death* remain taboo in U.S. society in favor of euphemisms such as *passing*, *going to a better place*, and so on. A failure to use explicit terms might hinder the effectiveness of communication between physicians and patients (Levy, 2001; Vincent, 1997). Recognizing this problem, Researcher Z wanted to know how often health care providers, patients, or family members used explicit terms versus euphemisms. Under what circumstances are these explicit terms used? Her research question was How are the terms *die*, *dying*, and *death* used in clinician-patient communication when discussing hospice care, and what alternative terms are used?

Researcher Z designed a sampling plan to maximize the diversity of the sample around demographic characteristics of both the clinician and the patient/family. Patient characteristics included gender, age, diagnosis, and ethnic background. Clinician characteristics included gender, discipline, and area of specialization. Two types of communication events with patients who had received a terminal diagnosis were sampled. One was discharge teaching for hospitalized patients who were being transferred to home hospice, inpatient hospice, or skilled nursing facilities for end-of-life (EOL) care. The other communication event was clinician-patient/family conferences in out- or inpatient settings to plan EOL care. Fifty separate communication events were sampled for 50 different clinicians and patient/family pairs. The research proposal was approved by institutional review boards before data collection. Informed consent was obtained from each participant. All clinician-patient conversations were audiotape-recorded and transcribed verbatim.

Data analysis started with computer-assisted searches for occurrences of the terms *die*, *death*, and *dying* in the transcripts. Word frequency counts for each of the three death-related terms in a transcript were calculated and compared to the total length of the communication event. Researcher Z also coded the identity of the speaker, such as physician, nurse, patient, or family member. Frequency counts by type of speaker were calculated and compared to the total number of terms coded.

Next, Researcher Z tried to identify alternative terms or expressions used instead of *death*, *die*, or *dying*. Occurrences of these terms were counted both as a total number and for each alternative term. Frequencies of euphemisms versus direct terms were compared for type of speaker, demographic characteristics of clinician, and demographic characteristics of patient within each communication event and across the total sample.

The major study findings described the occurrences of the three explicit terms used in clinician-patient communication as compared to euphemistic terms. Comparisons across type of speaker and characteristics of clinicians and patients were made. The discussion of this study focused on exploring possible explanations for differences in the use of explicit versus euphemistic terms when discussing EOL care for different groups and in different situations.

of content (Holsti, 1969). In this analysis, the focus is on discovering underlying meanings of the words or the content (Babbie, 1992; Catanzaro, 1988; Morse & Field, 1995). In Researcher Z's study, the initial part of the analysis technique, to count the frequency of *death*, *die*, and *dying* is more accurately viewed as a quantitative approach. However, Researcher Z went on to identify alternative terms for death and to examine the contexts within which direct versus euphemistic terms were used. Hence, Researcher Z used a summative approach to qualitative content analysis.

Researchers report using content analysis from this approach in studies that analyze manuscript types in a particular journal or specific content in textbooks. Examples include studies examining content related to EOL care in medical textbooks (Rabow, Hardie, Fair, & McPhee, 2000), EOL care in critical care nursing textbooks (Kirchhoff, Beckstrand, & Anumandla, 2003), palliative care in nursing textbooks (Ferrell, Virani, Grant, & Juarez, 2000), death and bereavement in nursing textbooks (Ferrell, Virani, Grant, & Borneman, 1999), and spirituality in

nursing textbooks (McEwen, 2004). These researchers started with counting the pages that covered specific topics followed by descriptions and interpretations of the content, including evaluating the quality of the content. Others have compared the results of a content analysis with other data collected within the same research project, such as comparing preferences for various types of television programming with socioeconomic indicators of participants (Krippendorff, 1980).

In a summative approach to qualitative content analysis, data analysis begins with searches for occurrences of the identified words by hand or by computer. Word frequency counts for each identified term are calculated, with source or speaker also identified. Researcher Z wanted to know the frequency of words that were used to refer to death but also to understand the underlying contexts for the use of explicit versus euphemistic terms. He or she illuminated the context of euphemistic versus explicit terms by reporting how their usage differed by variables such as the speaker (patient versus clinician), the clinician's specialization, and the age of the patient. Counting is used to identify patterns in the data and to contextualize the codes (Morgan, 1993). It allows for interpretation of the context associated with the use of the word or phrase. Researchers try to explore word usage or discover the range of meanings that a word can have in normal use.

A summative approach to qualitative content analysis has certain advantages. It is an unobtrusive and nonreactive way to study the phenomenon of interest (Babbie, 1992). It can provide basic insights into how words are actually used. However, the findings from this approach are limited by their inattention to the broader meanings present in the data. As evidence of trustworthiness, this type of study relies on credibility. A mechanism to demonstrate credibility or internal consistency is to show that the textual evidence is consistent with the interpretation (Weber, 1990). For Researcher Z's study, validation by content experts on what terms are used to replace the death terms would be essential. Alternatively, researchers can check with their participants as to their intended meaning through the process of member check (Lincoln & Guba, 1985).

SUMMARY OF KEY ASPECTS

All approaches to qualitative content analysis require a similar analytical process of seven classic steps, including formulating the research questions to be answered, selecting the sample to be analyzed, defining the categories to be applied, outlining the coding process and the coder training, implementing the coding process, determining trustworthiness, and analyzing the results of the coding process (Kaid, 1989). We have outlined how this process differs depending on the specific content analysis approach used. The success of a content analysis depends greatly on the coding process. The basic coding process in content analysis is to organize large quantities of text into much fewer content categories (Weber, 1990). Categories are patterns or themes that are directly expressed in the text or are derived from them through analysis. Then, relationships among categories are identified. In the coding process, researchers using content analysis create or develop a coding scheme to guide coders to make decisions in the analysis of content. A coding scheme is a translation device that organizes data into categories (Poole & Folger, 1981). A coding scheme includes the process and rules of data analysis that are systematic, logi-

TABLE 4: Major Coding Differences Among Three Approaches to Content Analysis

Type of Content Analysis	*Study Starts With*	*Timing of Defining Codes or Keywords*	*Source of Codes or Keywords*
Conventional content analysis	Observation	Codes are defined during data analysis	Codes are derived from data
Directed content analysis	Theory	Codes are defined before and during data analysis	Codes are derived from theory or relevant research findings
Summative content analysis	Keywords	Keywords are identified before and during data analysis	Keywords are derived from interest of researchers or review of literature

cal, and scientific. The development of a good coding scheme is central to trustworthiness in research using content analysis (Folger, Hewes, & Poole, 1984).

Key differences among conventional, directed, and summative approaches to content analysis center on how initial codes are developed. In a conventional content analysis, categories are derived from data during data analysis. The researcher is usually able to gain a richer understanding of a phenomenon with this approach. With a directed content analysis, the researcher uses existing theory or prior research to develop the initial coding scheme prior to beginning to analyze the data (Kyngas & Vanhanen, 1999). As analysis proceeds, additional codes are developed, and the initial coding scheme is revised and refined. Researchers employing a directed approach can efficiently extend or refine existing theory. The summative approach to content analysis is fundamentally different from the prior two approaches. Rather than analyzing the data as a whole, the text is often approached as single words or in relation to particular content. An analysis of the patterns leads to an interpretation of the contextual meaning of specific terms or content (Table 4).

CONCLUSIONS

Different research purposes require different research designs and analysis techniques (Knafl & Howard, 1984). The question of whether a study needs to use a conventional, directed, or summative approach to content analysis can be answered by matching the specific research purpose and the state of science in the area of interest with the appropriate analysis technique.

It is important for health researchers to delineate the specific approach to content analysis they are going to use in their studies before beginning data analysis. Creating and adhering to an analytic procedure or a coding scheme will increase trustworthiness or validity of the study. Careful description of the type of approach to content analysis used can provide a universal language for health researchers and strengthen the method's scientific base. Examples used in this article were drawn from the area of research on end of life, but the content analysis techniques described could be used in a broad range of studies. Content analysis offers researchers a flexible, pragmatic method for developing and extending knowledge of the human experience of health and illness.

REFERENCES

Babbie, E. (1992). *The practice of social research*. New York: Macmillan.

Barcus, F. E. (1959). *Communications content: Analysis of the research 1900-1958 (A content analysis of content analysis)*. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.

Berelson, B. (1952). *Content analysis in communication research*. Glencoe, IL: Free Press.

Budd, R. W., Thorp, R. K., & Donohew, L. (1967). *Content analysis of communications*. New York: Macmillan.

Catanzaro, M. (1988). Using qualitative analytical techniques. In N. F. Woods & M. Catanzaro (Eds.), *Nursing research: Theory and practice* (pp. 437-456). St. Louis, MO: C. V. Mosby.

Cavanagh, S. (1997). Content analysis: concepts, methods and applications. *Nurse Researcher*, *4*(3), 5-16.

Coffey, A., & Atkinson, P. (1996). *Making sense of qualitative data: Complementary research strategies*. Thousand Oaks: Sage.

Curtis, J. R., Wenrich, M. D., Carline, J. D., Shannon, S. E., Ambrozy, D. M., & Ramsey, P. G. (2001). Understanding physicians' skills at providing end-of-life care: Perspectives of patients, families, and health care workers. *Journal of General Internal Medicine*, *16*, 41-49.

Downe-Wamboldt, B. (1992). Content analysis: Method, applications, and issues. *Health Care for Women International*, *13*, 313-321.

Ferrell, B., Virani, R., Grant, M., & Borneman, T. (1999). Analysis of content regarding death and bereavement in nursing texts. *Psychooncology*, *8*, 500-510.

Ferrell, B., Virani, R., Grant, M., & Juarez, G. (2000). Analysis of palliative care content in nursing textbooks. *Journal of Palliative Care*, *16*, 39-47.

Folger, J. P., Hewes, D. E., & Poole, M. S. (1984). Coding social interaction. In B. Dervin & M. J. Voigt (Eds.), *Progress in communication sciences* (pp. 115-161). Norwood, NJ: Ablex.

Hickey, G., & Kipping, C. (1996). Issues in research. A multi-stage approach to the coding of data from open-ended questions. *Nurse Researcher*, *4*, 81-91.

Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley.

Kaid, L. L. (1989). Content analysis. In P. Emmert & L. L. Barker (Eds.), *Measurement of communication behavior* (pp. 197-217). New York: Longman.

Kirchhoff, K. T., Beckstrand, R. L., & Anumandla, P. R. (2003). Analysis of end-of-life content in critical care nursing textbooks. *Journal of Professional Nursing*, *19*, 372-381.

Knafl, K. A., & Howard, M. J. (1984). Interpreting and reporting qualitative research. *Research in Nursing and Health*, *7*, 17-24.

Kondracki, N. L., & Wellman, N. S. (2002). Content analysis: Review of methods and their applications in nutrition education. *Journal of Nutrition Education and Behavior*, *34*, 224-230.

Krippendorf, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills, CA: Sage.

Kübler-Ross, E. (1969). *On death and dying*. New York: Macmillan.

Kyngas, H., & Vanhanen, L. (1999). Content analysis as a research method [Finnish]. *Hoitotiede*, *11*, 3-12.

Levy, M. M. (2001). End-of-life care in the intensive care unit: Can we do better? *Critical Care Medicine*, *29*, N56-N61.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Lindkvist, K. (1981). Approaches to textual analysis. In K. E. Rosengren (Ed.), *Advances in content analysis* (pp. 23-41). Beverly Hills, CA: Sage.

Manning, K. (1997). Authenticity in constructivist inquiry: Methodological considerations without prescription. *Qualitative Inquiry*, *3*, 93-115.

Mayring, P. (2000). Qualitative content analysis. *Forum: Qualitative Social Research*, *1*(2). Retrieved March 10, 2005, from http://www.qualitative-research.net/fqs-texte/2-00/02-00mayring-e.htm

McEwen, M. (2004). Analysis of spirituality content in nursing textbooks. *Journal of Nursing Education*, *43*, 20-30.

McTavish, D.-G., & Pirro, E.-B. (1990). Contextual content analysis. *Quality and Quantity*, *24*, 245-265.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage.

Morgan, D. L. (1993). Qualitative content analysis: A guide to paths not taken. *Qualitative Health Research*, *3*, 112-121.

Morse, J. M. (1991). *Qualitative nursing research*. Newbury Park, CA: Sage.

Morse, J. M., & Field, P. A. (1995). *Qualitative research methods for health professionals* (2nd ed.). Thousand Oaks, CA: Sage.

Nandy, B. R., & Sarvela, P. D. (1997). Content analysis reexamined: A relevant research method for health education. *American Journal of Health Behavior*, *21*, 222-234.
National Institute of Nursing Research. (2003). *Research themes for the future*. Retrieved October 10, 2003, from http://www.nih.gov/ninr/research/themes.doc
Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.
Poole, M. S., & Folger, J. P. (1981). Modes of observation and the validation of interaction analysis schemes. *Small Group Behavior*, *12*, 477-493.
Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research*, *27*, 258-284.
Rabow, M. W., Hardie, G. E., Fair, J. M., & McPhee, S. J. (2000). End-of-life care content in 50 textbooks from multiple specialties. *Journal of the American Medical Association*, *283*, 771-778.
Rosengren, K. E. (1981). Advances in Scandinavia content analysis: An introduction. In K. E. Rosengren (Ed.), *Advances in content analysis* (pp. 9-19). Beverly Hills, CA: Sage.
Tesch, R. (1990). *Qualitative research: Analysis types and software tools*. Bristol, PA: Falmer.
Vincent, J.-L. (1997). Communication in the ICU. *Intensive Care Medicine*, *23*, 1093-1098.
Weber, R. P. (1990). *Basic content analysis*. Beverly Hills, CA: Sage.
Wilson, C. T., & Fletcher, P. C. (2002). Dealing with colon cancer: One woman's emotional journey. *Clinical Nurse Specialist*, *16*, 298-305.

Hsiu-Fang Hsieh, Ph.D., is an assistant professor at Fooyin University, Kaohsiung Hsien, Taiwan.

Sarah E. Shannon, Ph.D., is an associate professor at the University of Washington, Seattle.