

## Chapter 3

# Content Analysis

Content analysis consists of procedures for defining, measuring, and analyzing both the substance and meaning of texts or messages or documents. Over the years, the definition of content analysis has evolved to embrace larger contexts. Bernard Berelson, an early pioneer of the methodology in social science research, defined content analysis as a “technique for the objective, systematic, and quantitative description of manifest content of communication” (1952, 74). Writing a half-century later, Kimberly Neuendorf (2002, 10) described the methodology as primarily quantitative:

Content analysis is a summarizing, quantitative analysis of messages that relies on the scientific method (including attention to objectivity-intersubjectivity, a priori design, reliability, validity, generalizability, replicability, and hypothesis testing) and is not limited as to the types of variables that may be measured or the context in which the messages are created or presented.

It is important to keep in mind, however, that content analysis can be either quantitative or qualitative. When researchers take a quantitative approach, they focus on numerically measurable objectives. Their research questions are typically stated as hypotheses, they use standardized instruments of proven reliability and validity, and they use inferential statistical techniques in data analysis. Although Neuendorf firmly believes that this methodology can only be quantitative, not all content analysis research concerns itself with the counting of things. Klaus Krippendorff (2004, 87) points out that “quantification is not a defining criterion for content analysis.” In fact, he notes that “[u]sing numbers instead of verbal categories or counting instead of listing quotes is merely convenient; it is not a requirement for obtaining valid answers to a research question.” Both quantitative and qualitative approaches to content analysis are valid and coexist within the social sciences and within library and information science research. They are simply different ways of examining the same problem. We will look at both quantitative and qualitative content analysis studies within this chapter.

Content analysis has been used in social science research for the better part of a century as a means to determine message characteristics found in a body of text. In fact, some type of content analysis has been applied to the study of texts for several centuries. Then and now, most content analysis research has focused on the news media. In the seventeenth century, the Catholic church conducted some of the earliest content analysis studies because they feared that newspapers—then a relatively new communication medium—were spreading irreligious information. In the nineteenth century, with the dramatic increase in mass-produced newspapers, many social critics conducted content analyses on the types and the tone of newspaper articles. Some studies measured the number of column inches devoted to religious, scientific, and literary topics, as opposed to stories dealing with sports, gossip, or scandal; other studies focused on the numbers of significant and wholesome stories versus those that were deemed to be cheap, frivolous, or immoral. Although early efforts in content analysis may now seem simplistic and decidedly biased, content analysis has evolved as a research methodology over the past century to become rigorous, grounded in theory, and standardized in its procedures. Throughout the twentieth century, the methodology experienced an intellectual growth as social scientists expanded their studies to different types of texts and began asking new kinds of questions. In the mid-to latter part of the last century, the methodology spread to other disciplines such as psychology, linguistics, sociology and, not surprisingly, library science.

Beginning in the 1940s and continuing for the next forty or so years, volumes of *Library Literature* indicate content analysis as a popular methodology for thesis and dissertation research. Bernard Berelson, an early pioneer in the development of content analysis as a research methodology, served as dean of the Library School at the University of Chicago from 1946 to 1951. It was primarily due to his influence, as well as overall trends in social science research during the mid-part of the twentieth century, that library science began using content analysis as a research methodology. From 1943 to 1963, 62 percent of all theses and dissertations listed in *Library Literature* used content analysis. One of Berelson's students at the University of Chicago was Lester Asheim, whose dissertation compared the content of novels and the motion pictures made from them, and conducted many other research projects that dealt with themes or social issues found in specific media (book reviews, popular magazines, children's and young adult literature).

The methodology is still one of the more popular ones within the field. In a 2004 study on research in library science, Denise Koufogiannakis and Linda Slater found that content analysis is one of the top five preferred methodologies in library science literature. (For a broad historical overview of content analysis, see Krippendorf [2004, 3–17] and Berelson [1952, 21–25], which provide a brief overview of early twentieth-century content analysis research conducted in the United States.)

In conducting a content analysis study, a text's content is revealed by looking at it in a manner different from the ordinary reading of the text. Note also that the term "text" is considered in its broadest sense. Very loosely defined, a "text" consists of any material type that communicates meaning. Consider the range of texts that can be studied using content analysis:

- *Written materials:* books, journals, magazines, newspapers, advertisements, official documents.
- *Visual items:* films, documentaries, television programs or advertisements, photographs, works of art, clothing.
- *Sound texts:* music lyrics, operas, musicals, songs, polkas.
- *Combinations of types of materials:* Web pages, performance art, or computer programs that combine visual, text, and sound elements.

The "content" of a given text is anything (words, phrases, pictures, ideas) that can be communicated within the text. To illustrate how the content is drawn out of and studied within a body of texts, take for example a hypothetical study of how Hollywood has depicted dog intelligence in feature films over the past 75 years. In this case, the "texts" might consist of all Hollywood-produced feature films with at least one dog in a main or supporting role. The content studied could be made up of actions, signals, behaviors or any other communication modes whereby canines have exhibited intelligence through their actions. Generally speaking, after defining and categorizing these "intelligence signs," researchers would then review all the films sampled, code them, and then compare content across films (the texts) to find similarities, themes, or trends. Researchers could analyze and comment on their findings either through reporting raw numbers and percentages or by pulling examples or quotes from the texts to illustrate main points. Another example of "texts" and "content" is Kuchi's (2006) study examining how and to whom academic libraries communicate their mission statements in the Web environment. The "texts" Kuchi studied were the library mission statements of the 111 Web-based Association of Research Libraries (ARL), whereas the "content" she reviewed consisted of the visibility of these mission statements (statements linked directly from the home page or linked indirectly from other pages) and the mission statements' intended audience (links to the mission statements from specific stakeholders' Web pages, e.g., students).

Drawbacks to conducting content analysis research are relatively few, but should be noted. In practice, content analysis can be time-consuming and labor intensive. Coding documents by hand requires time, patience and, often, assistance. Even if the researcher uses computer programs to count occurrences and frequencies, the computer program is only as good as the programmer who creates it and the person who uses it. Users of computer programs for analysis still must create and employ well-defined categories for coding content. Content analysis also cannot be used to make claims about motives, the meanings that

individuals draw from messages, or the effects of those messages. It can reveal trends and themes as evidenced through the texts studied, but it does not show the cause or the result of those trends, themes or behaviors. Finally, and perhaps most importantly, "content analysis cannot determine the truthfulness of an assertion... it reveals the content in a text but it cannot interpret the content's significance" (Neuman 2003, 311).

Nonetheless, content analysis is a very attractive method for library researchers because it is nonreactive, unobtrusive, and not limited by geography. It is nonreactive in the sense that people themselves are not being studied; rather, textual evidence of their social behavior or actions is examined. It is unobtrusive in the sense that researchers are not directly studying human behavior; instead, the method looks at human-produced artifacts. This means that content analysis studies are often exempt from review by institutional review boards, or are subjected to more lenient review than studies using human subjects. (In fact, content analysis raises few, if any, ethical concerns, especially when the study uses publicly available texts.) The unobtrusiveness of content analysis also means that content analyses are free from the reactivity effects that human subjects often introduce into research studies; newspaper articles, after all, cannot shape their "responses" to match researchers' presumed intentions like people can.

Content analysis is not limited by time or space. Content analysis can be used to study the past through analysis of diaries, newspapers, or other archival-type records. In this way, it serves as an important learning tool, giving a peephole view on the concerns of people from past centuries. Because it is not confined to a specific area, researchers can just as easily study content produced on other continents as they can study content produced locally. Moreover, depending on the type of text under examination, there are few history or maturation effects even when the study is interrupted. Texts' unchanging format, coupled with the fact that most texts are portable, is very handy for researchers, who can then study the texts at any time and in any place. If Web sites are the object of examination, the researcher needs to take measures to ensure a tight data collection period because of their volatility. Web sites aside, most texts under investigation are in a fixed format and are not likely to change.

## FINDING A TOPIC

Content analysis studies fall into three different broad categories: text-driven, method-driven, and problem-driven (Krippendorf 2004). Text-driven analyses arise out of the text itself and often begin without a specific research question in mind. They are exploratory in nature, seeking to arrive at a general understanding of the texts, or collection of texts, under analysis. And because they are exploratory, not setting up hypotheses to test or measures to apply, they are often qualitative in nature. Very few content analysis studies in library/science are text-driven.

Method-driven analyses come from the researcher's wish to apply the methodology to previously unexplored areas. An example of a method-driven analysis would be taking a methodology developed to analyze gender roles in advertisements and applying it to ads appearing in library science publications.

A problem-driven analysis involves real world problems that reflect the concerns and issues within a discipline or a focus area. Epistemic in nature, problem-driven content analysis studies purposefully examine texts in the belief that a methodical examination of these will provide answers to research questions. Almost all content analysis research in library and information science within the past 20 years has been problem-driven, as is indicated by the fact that all the studies described in Figure 3-1 or mentioned in the following categories are problem-driven.

**Figure 3-1: Studies Using Content Analysis Methodology**

Dahl, Candice. 2001. "Electronic Pathfinders in Academic Libraries: An Analysis of Their Content and Form." *College & Research Libraries* 62 (3): 227-237.

*Using guidelines developed for print pathfinders (Kapoun 1995), the researcher analyzed the content of 45 electronic pathfinders from 9 academic institutions in Canada. She examined the pathfinders for consistency of format within the institution, scope, readability, and usability, ranking the pathfinders in each of the 4 categories on a 1 to 3 scale. Findings show little uniformity among pathfinders within each institution and that pathfinders vary in their complexity and usefulness.*

Ellison, Jim. 2004. "Assessing the Accessibility of Fifty United States Government Web Pages: Using Bobby to Check on Uncle Sam." *First Monday* 9 (7). Available: [http://firstmonday.org/issues/issue9\\_7/ellison/index.html](http://firstmonday.org/issues/issue9_7/ellison/index.html)

*As a way of determining whether all 50 Web sites met federally-required accessibility guidelines, the researcher used two different computer programs, Bobby and Cynthia Says, to evaluate the White House, 2 government Web portals, Web sites from 15 executive branch departments and 32 federal agencies. Provides descriptive chart listing each government Web site, its url and type and severity of accessibility violations.*

Hahn, Karla L., and Kari Schmidt. 2005. "Web Communications and Collections Outreach to Faculty." *College & Research Libraries* 66 (1): 28-40.

*Researchers looked at 149 SPARC member Web sites on collection development to see how libraries are communicating changes in their collections and issues in scholarly communication. They created a pilot survey to develop a coding scheme for the 4 categories used to examine libraries' collection development Web. Findings are reported in raw numbers and percentages, and researchers include charts showing responses to questions 1 through 3.*

Koufogiannakis, Denise, Linda Slater, and Ellen Crumley. 2004. "A Content Analysis of Librarianship Research." *Journal of Information Science* 30 (3): 227-239.

*Researchers examined 2,664 articles published in 91 library science journals in 2001 as an effort to determine the prevalence of research within the profession's journal literature, the characteristics of the type of research conducted as well as the topics covered. Researchers performed independent assessments with high inter-rater reliability and found that a little over 30 percent of articles published were research. They list the top library science research journals, also charting study types by domain and the top 5 research journals by domain.*

(Cont'd.)

**Figure 3-1: Studies Using Content Analysis Methodology (Continued)**

McGrath, Eillen L., Winifred Fordham Metz, and John B. Rutledge. 2005. "H-Net Book Reviews: Enhancing Scholarly Communication with Technology." *College & Research Libraries* 66: 8-19.

*Authors studied the differences between traditional and electronic book reviews by examining the length, content, style, timeliness, and format of book reviews appearing on H-Net in 2002 and found significant differences in length and review style of electronic reviews.*

Snelson, Pamela, and S. Anita Talar. 1991. "Content Analysis of ACRL Conference Papers." *College & Research Libraries* 52 (3): 466-472.

*Authors examined the content of 181 papers presented at the 2nd through 4th ACRL conferences to determine whether papers were more research based than found in a former study where only one-third of the papers were research. Descriptive tables note the overall research content from each conference, major research goals, and research characteristics.*

Still, Julie. 1998. "Role and Image of the Library and Librarians in Discipline-Specific Pedagogical Journals." *Journal of Academic Librarianship* 24 (3): 225-231.

*The author searched ERIC and conducted a page-by-page review of 13 of the journals to characterize and describe the image and role of librarians in 19 discipline-specific teaching journals. In the seven-year period under examination, 1990-1996, the author found that only 53 articles from the 13,016 listed in the 29 journals mentioned libraries or librarians. Excerpts from articles mentioning libraries are provided showing how the institution and the profession are portrayed within the literature.*

Further, content analysis studies tend to focus on one of three different topical areas:

- *Studies that are focused on the delivery of library services.* These range from the characteristics of interactive reference service via the Web (Bao 2003; Wells 2003), to commonalities in teaching Web-based full-text databases (Bernard and Hollingsworth 1999), to online tutorials (Dewald 1999; Tancheva 2003), to materials for presenting scholarly communication and collection development information to academic faculty (Hahn and Schmidt, as listed in Figure 3-1). In service delivery studies, researchers are typically interested in identifying similarities, characteristics and themes in texts in order to improve or enhance the particular services within their own library that are embodied by those texts. For instance, in the Hahn and Schmidt study, the authors wished to discover how their collection management pages could be improved "to promote awareness of recent collection changes" (2005, 29). Similarly, Kornelia Tancheva (2003) examined library instruction tutorials to see how they adopted learning theory principles while she was working on designing an online tutorial for Cornell University's Mann Library.
- *Studies that are focused on specific resources commonly found in libraries.* Often viewed as the prototypical content analysis study in library science, resource-specific studies examine a particular type of work as a way to

learn more about how the resource or genre functions and to discover similarities with other types of resources. As shown in Figure 3-1, McGrath, Metz and Rutledge's research on electronic book reviews is a good example of a resource-specific study. They compared book reviews appearing on the electronic distribution list H-Net over the course of one year with print book reviews in order to determine how electronic book reviews differ from those in traditional sources. Book reviews are frequent subjects for content analysis studies, as are newspapers, especially online newspapers, which have gained a lot of attention within the past few years because of the innovative electronic functions they make available to users. Erdelez and Rioux (2000) looked at the different ways online newspapers allow readers to share articles with others ("e-mail this article," "share this article with a friend") and Doughty (2002) evaluated ten online newspapers using standard usability guidelines. Web pages are another frequent subject of content analyses. Ryan, Field and Olfman (2003) focused on state government home pages, looking at how these changed over a five-year period, while Julie Still (2001) studied university library Web sites from English-speaking countries.

- *Studies that are focused on the profession itself.* Like members of other professions, library and information scientists are concerned with the evolution and changing nature of their work and their workplace. Of particular concern is the enormous impact that technology has had in libraries. Nowhere is this concern more apparent than in the profusion of job-trend studies using content analysis over the past 15 to 20 years (listed in Figure 3-2). Studies such as these benefit the profession as a whole because they not only can serve as a bellwether for changes within the profession but also signal possible changes in budget and resource allocations. These profession-focused studies also provide proof for those needing a scholarly article to point to as they make their case for an issue. In fact, critics of library school curricula, as well as instruction librarians, have commonly pointed to Lynch and Smith's (2001) study of academic librarian positions to bolster their claims that library school offerings do not correspond to workplace needs for computer and teaching proficiencies.

Although "[n]o single model is applicable to all content analyses" (Busha and Harter 1980, 174), all content analysis studies do share common elements. Pamela Snelson and Anita Talar's 1991 article analyzing ACRL conference papers for research components serves as a good illustration of these common elements. Nonreactive in its approach, Snelson and Talar's study analyzed all conference papers presented at the second through fourth ACRL conferences to determine whether or not "true" research was increasing within the profession. Each of the 181 papers analyzed served as a unit of analysis for the study. Because the total number of papers, or units, was relatively small, the authors did not need to

Figure 3-2: Job Analysis Studies

Job Type	Description of study	Citation
Academic librarian	The authors updated Reser and Schuneman's prior work and tracked changes in the academic job market by examining 900 job ads published in 1996 from <i>American Libraries</i> , <i>The Chronicle of Higher Education</i> , <i>College &amp; Research Libraries</i> , and <i>Library Journal</i> . Authors specifically examined differences in public, systems, and technical services positions.	Belle, Penny, and Megan M. Adams. 2000. "Other Duties as Assigned: Emerging Trends in the Academic Library Job Market." <i>College &amp; Research Libraries</i> 61: 336-347.
Academic librarian	Authors culled 220 job ads from the March issues of <i>College &amp; Research Libraries News</i> every 5th year from 1973 to 1998 to discover how job requirements have changed over time. Authors specifically examined behavioral skills required, degree requirements and faculty status, noting emerging trends in the need for computer and teaching skills.	Lynch, Beverly P., and Kimberley Robles Smith. 2001. "The Changing Nature of Work in Academic Libraries." <i>College &amp; Research Libraries</i> 62: 407-420.
Academic librarian	Authors examined the differences in public and technical services positions found in 1,133 job ads published in 1988 from <i>American Libraries</i> , <i>College &amp; Research Libraries</i> , and <i>Library Journal</i> and looked in particular at job skills, degree and foreign language requirements, and salary ranges.	Reser, David, and Anita Schuneman. 1992. "The Academic Library Job Market: A Content Analysis Comparing Public and Technical Services." <i>College &amp; Research Libraries</i> 53: 49-59.
Academic librarian	Specifically addressing computer skills, the author reviewed 2,500 job ads posted every 5 years in <i>American Libraries</i> over a 20-year period.	Zhou, Yuan. 1996. "Analysis of Trends in Demand for Computer-Related Skills for Academic Librarians from 1974 to 1994." <i>College &amp; Research Libraries</i> 57: 259-272.
Cataloger	In determining technology's impact on cataloging positions, the author reviewed 151 job ads appearing in <i>American Libraries</i> and <i>College &amp; Research Libraries News</i> from 2000 to 2002 and specifically focused on position title, degree requirements, and necessary skills.	Khurshid, Zahiruddin. 2003. "The Impact of Information Technology on Job Requirements and Qualifications for Catalogers." <i>Information Technology and Libraries</i> 22: 18-21.
Cataloger	Studied changes in cataloging positions over a ten-year period by culling job postings in one issue per year in <i>American Libraries</i> .	Chaudhry, Abdus Sattar, and N. C. Komathi. 2001. "Requirements for Cataloging Positions in the Electronic Environment." <i>Technical Services Quarterly</i> 19: 1-23.

(Cont'd.)

Figure 3-2: Job Analysis Studies (Continued)

Job Type	Description of study	Citation
Collection development librarian	Researcher drew on 433 job ads posted in <i>College &amp; Research Libraries News</i> over 11 years to determine what type of skills and experience were required for collection development positions.	Robinson, William C. 1993. "Academic Library Collection Development and Management Positions: Announcements in <i>College &amp; Research Libraries News</i> from 1980 to 1991." <i>Library Resources &amp; Technical Services</i> 37: 134-146.
Electronic resources librarian	Author traced the evolution of the position over a 17-year period, from 1985 to 2001, through an examination of 298 jobs posted in <i>American Libraries</i> .	Fisher, William. 2003. "The Electronic Resources Librarian Position: A Public Services Phenomenon?" <i>Library Collections, Acquisitions &amp; Technical Services</i> 27: 3-17.
Electronic resources/digital librarian	Researchers examined 223 job ads appearing in <i>College &amp; Research Libraries News</i> from 1990 to 2000, focusing specifically on job title, reporting line, duties and home department, in order to learn about the evolving nature of technology in academic libraries.	Cronels, Karen S., and Pat Henderson. 2002. "Electronic and Digital Librarian Positions: A Content Analysis of Announcements from 1990 through 2000." <i>Journal of Academic Librarianship</i> 28: 232-237.
Instruction librarian	Author used both qualitative and quantitative methods in examining LI positions posted over three-month period on LIBJOBS to determine skills and knowledge required.	Clyde, Laurel A. 2002. "An Instructional Role for Librarians: An Overview and Content Analysis of Job Advertisements." <i>Australian Academic &amp; Research Libraries</i> 33: 150-166.
Preservation librarian	Authors tracked the evolution of preservation librarian positions by examining job titles, requirements, education and experience found in 116 preservation-type positions posted in five different publications over a 13-year period.	Cloonan, Michele Valerie, and Patricia C. Norcott. 1989. "Evolution of Preservation Librarianship as Reflected in Job Descriptions from 1975 to 1987." <i>College &amp; Research Libraries</i> 50: 646-656.
Subject specialist librarian	Author studied 315 job announcements appearing in three different publications over eight and one-half years to see how the position has evolved and specifically looked at position title, ARL status of hiring institution, tenure requirements, salary range, duties, reporting line, and both required and desired education, skills, and experience.	White, Gary W. 1999. "Academic Subject Specialist Positions in the United States: A Content Analysis of Announcements from 1990 through 1998." <i>The Journal of Academic Librarianship</i> 25: 372-382.

(Cont'd.)

Figure 3-2: Job Analysis Studies (Continued)

Job Type	Description of study	Citation
System librarian	Author uses qualitative techniques to determine characteristics of systems librarian positions (job title, duties, skills, degree requirements, reporting line, salary) through an analysis of 107 announcements posted in <i>College &amp; Research Libraries News</i> over a four-year period.	Footo, Margaret. 1997. "The Systems Librarian in U.S. Academic Libraries: A Survey of Announcements from <i>College &amp; Research Libraries News</i> , 1990-1994." <i>College &amp; Research Libraries</i> 58: 517-526.
Youth services librarian	In order to determine whether demand has increased for youth services librarians over time and whether changes have occurred in employment criteria and job duties, Adkins examined job titles, responsibilities, education, skills, and experience requirements found in 285 youth services positions posted in <i>American Libraries</i> from 1971 to 2001 in five-year increments.	Adkins, Denise. 2004. "Changes in Public Library Youth Services: A Content Analysis of Youth Services Job Advertisements." <i>Public Library Quarterly</i> 23: 59-73.

sample. They had already narrowed down their population by examining only three sets of conference papers. They were also fortunate in that a previous study by Coughlin and Snelson (1983) on a similar topic had already established coding categories and corresponding conceptual definitions that they could use. (As further proof that research can proceed by accretion, Coughlin and Snelson themselves drew on conceptual definitions and categories previously developed by Atherton (1975) in her study of the research methods employed in information science literature.) These "recycled" conceptual definitions provided the criteria by which Snelson and Talar judged whether or not the papers were research reports. They also furnished Snelson and Talar with previously validated categories by which to structure their analysis of the specific characteristics common to all of the units studied. Snelson and Talar tested inter-coder reliability, also referred to as inter-rater reliability, in a small pilot study where each researcher coded a sample of conference papers and achieved a .90 reliability coefficient. The variables in this study (the categories studied) were nominal, or non-numeric, categorical variables that cannot be counted (i.e., gender, race, religious affiliation are all nominal variables). Because they used nominal variables, Snelson and Talar used chi-square ( $\chi^2$ ) analysis as an inferential statistic to accept the null hypothesis of no significant increase in the number of research papers (chi-square and null hypothesis will be discussed in more detail later in the text).

However, as mentioned earlier, not all content analysis studies are quantitative, as Snelson and Talar's was; some studies are qualitative in nature. A good example of a qualitative content analysis study is Foote's analysis of job postings

for systems librarians (listed in Figure 3-2). She took an inductive approach in identifying and analyzing job requirement categories for systems librarians. Some of these are not visible at the outset of the study but emerge as important concepts through careful analysis. Foote's study also features other important characteristics found in qualitative studies: a relatively small sample size, no statistical measures employed, and findings reported in raw numbers and quotations for the text. For a more in-depth description of the differences between quantitative and qualitative research see Chapter 4.

### FORMULATING QUESTIONS

Serving as a guide to the overall research design, the well developed research question provides the infrastructure for the entire research project. Content analysts typically formulate research questions that can have several possible solutions and that deal with previously unobserved phenomena within the text studied. Like many other methodologies, a content analysis study frequently presents one or more hypotheses in addition to the research question. A hypothesis differs from a research question. A hypothesis is a declarative statement that predicts a relationship between two or more variables, while a research question asks about an observed reality. Consider these research questions and corresponding hypotheses drawn from content analysis research in library and information science:

- *Research question:* "Have requirements for entry level positions become more stringent or lax over time?"  
*Hypothesis:* "Over time, employers require more experience and knowledge that cannot always be gained from library school" (Sproles and Ralledge 2004, under "Methodology").
- *Research question:* To what degree has the U.S. government met its goal of rendering its Web pages accessible to people with disabilities?  
*Hypothesis:* The United States government has met Section 508 accessibility guidelines for all of its Web pages (Ellison 2004, Figure 3-1).
- *Research question:* Which job announcements are more likely to require that candidates possess an advanced subject degree, technical services or public services librarian positions?  
*Hypothesis:* "Public services jobs are more likely to require advanced subject degrees" (Beile and Adams 2000, 337, Figure 3-2).
- Research question: Do library and information science research papers differ from their science counterparts?  
*Hypothesis:* Library science conference papers differ from their science counterparts by being less likely to include a problem statement, literature review, hypothesis, research methodology, findings and conclusions. They are, therefore, less rigorous. (Snelson and Talar 1991, 468, Figure 3-1).

Not all content analysis research in library and information science contains formal hypotheses; however, researchers should have a hypothesis for such a project before conducting it even if they do not state it, or state it only informally. Hypotheses, whether they are stated or unstated, formal or informal, help guide the researchers' reading of the texts. Similarly, even where content analyses have hypotheses, they may report their data in raw numbers and percentages without employing tests of statistical significance to test these hypotheses. Such is the case in Cloonan and Northcott's (1989) article on the evolving role of the preservation librarian, as well as in Laurel Clyde's review of instruction librarian positions. Both studies deal with small data sets and both explore the subtle, but significant, changes in two different librarian positions over time.

An important element in developing research questions and hypotheses is the formulation of conceptual definitions for the variables within the study. Without solid, well-thought-out descriptions of categories and carefully devised coding procedures, content analysis is weak. These categories are classes of characteristics, topics, or themes that recur regularly enough to be quantified and described. Content analysts either rely on standard definitions or create their own definitions for categories. Wherever possible, researchers try to use standard definitions to increase the validity of their own research and heighten its reliability. Koufogiannakis, Slater and Crumley (2004), for example, employed a previously developed taxonomy of six separate subject domains (reference, education, collections, management, information access, retrieval and marketing) to classify the topics in library science research articles published in 2001 (described in Figure 3-1). However, where standard definitions do not exist, researchers must construct their own. For example, in studying how academic libraries communicate their mission statements via the Web, Kuchi (2006) created definitions and categories for direct links from the library main page; indirect links from the library main page; and the labels used to identify the link to the mission statement. Regardless of whether researchers create their own definitions for categories or rely on standard definitions, it is essential that the definitions be exhaustive and mutually exclusive. This is because the definitions form the basis for the codes, which will be later applied to the data set. If clearly defined categories are absent in a study, the project loses its conceptual framework and, essentially, loses its *raison d'être*. For similar reasons, good definitions are required for the researcher to implement an objective and systematic counting-and-recording procedure that produces "a quantitative description of the symbolic content in a text" (Neuman 2003, 211).

### **DEFINING THE POPULATION**

Population definition actually occurs at several steps within a content analysis research project. It is sometimes inherent in the topic selection mechanism

when the study is purposely limited by date, by type of text, or by group within the research question. Dahl (2001) uses population definition as a way of simultaneously selecting a topic by focusing on electronic pathfinders from nine Canadian libraries (listed in Figure 3-1). By looking only at *electronic* pathfinders, she has narrowed down the entire universe of pathfinders to focus on those appearing in a specific format. Even more specifically, Dahl looks at Web-based pathfinders—not all electronic pathfinders. She further defines her population to include pathfinders from "the library web sites of top ranking Canadian universities according to the annual ranking conducted by *Maclean's* in 1999" (Dahl 2001, 229). She thus limits her focus by geographic region as well as institution type. (Moreover, her use of a well-recognized external source such as *Maclean's* helps to ensure the validity of her project.) Similarly, Hahn and Schmidt in their analysis of faculty outreach and communication on collection development Web pages narrowed their population early on by asking their research question only of Scholarly Publishing and Academic Resources Coalition (SPARC) member Web sites. (SPARC was developed by the Association of Research Libraries and is "an alliance of universities, research libraries, and organizations" whose purpose is to serve as a "constructive response to market dysfunctions in the scholarly communication system." [www.arl.org/sparc/about/index.html](http://www.arl.org/sparc/about/index.html).) They singled out SPARC member institutions, in part because SPARC members are at the forefront of addressing major issues in scholarly communication but also because they recognized that this group was small enough (149 members at the time of the study) to comprise a workable dataset.

However, not all content analysis studies in library science simultaneously define their populations at the research question development phase. Ellison's (2004) study, for example, asked how well U.S. government Web sites complied with Section 508 accessibility guidelines. Even with this question (which was limited to U.S. government Web sites as opposed to all Web sites), Ellison still needed to further define the population because of the high number of U.S. government sites. He ultimately used purposive sampling (see below) to select 50 home pages—chosen to reflect a balance between executive departments—from the approximately 20,000 U.S. government home pages (2004, Under "Methodology").

Even when topic selection does not also determine population selection, the topic under examination guides the library researcher in limiting the population. The job trend studies listed in Figure 3-2 limit their populations by position type as well as by time period and by publication. Zhou (1996) pulled job ads appearing only in *American Libraries* over a 20-year span; Clyde (2002) limited her analysis to job ads posted on the electronic distribution list, LIBJOBS, for three months; Belle and Adams (2000) selected a year's accumulation of academic library job ads published in four major professional journals. One might wonder why these authors took three separate approaches when