Library Trends

TRENDS

LIBRARY

Research Methods

Lynda M. Baker, Issue Editor

Published by The Johns Hopkins University Press

Volume 55 | Number 1 | Summer 2006

Volume 55
Number 1
Summer 2006

LT-55-1-cover.indd 1
8/14/06 1:54:45 PM

TRENDS

LIBRARY

Library Trends, a quarterly thematic journal, focuses on current trends in all areas of library practice. Each issue addresses a single theme in depth, exploring topics of interest primarily to practicing librarians and information scientists and secondarily to educators and students.

Editors: Boyd Rayward and John Unsworth
Publications Committee: Boyd Rayward, John Unsworth, Terry Weech, Marlo Welshons

© 2006 by The Board of Trustees of the University of Illinois. All rights reserved.

Library Trends (ISSN 0024-2594) is published quarterly—in Summer, Fall, Winter, and Spring—for the University of Illinois Graduate School of Library and Information Science (GSLIS) by The Johns Hopkins University Press, 2715 N. Charles Street, Baltimore, MD 21218-4363.

Permissions. No portion of *Library Trends* may be reproduced by any process or technique without the formal consent of the publisher. Copies for personal or internal use may be made on the condition that the copier pay a fee of $8.00 per copy through the Copyright Clearance Center, 222 Rosewood Dr., Danvers, MA 01970, for copying beyond that permitted by Section 107 or 108 of the U.S. Copyright Law. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. 0024-2594/05 $8.00. Direct all other permissions requests to Permissions Manager, The Johns Hopkins University Press, 2715 N. Charles Street, Baltimore, MD 21218-4363; or visit www.press.jhu.edu/cgi-bin/permissions.cgi.

Subscriptions. Institutional rate per year is $120; individual rate is $75.00. Students with ID may subscribe for $45.00. *For all subscriptions mailed outside the U.S.A., please add foreign shipping:* $7.00 (Canada or Mexico); $14.60 (outside North America). Single copy rate (current or back issues) is $22.50 for institutions; $22.50 for individuals. If applicable, please add state taxes or GST. *Prepayment is required for shipment.* All orders, address changes, and other business correspondence should be addressed to The Johns Hopkins University Press, Journals Division, P.O. Box 19966, Baltimore, MD 21211-0966 (USA); telephone (410) 516-6987; fax (410) 516-6968; email jrnlcirc@press.jhu.edu; or toll free 1-800-548-1784 (US and Canada only). Claims for replacement of missing issues must be received within three months of mailing (six months for foreign addresses). All notices of change of address should provide both the old and new address.

Periodicals postage paid at Baltimore, Maryland, and at additional mailing offices. POSTMASTER: Send address changes to *Library Trends,* The Johns Hopkins University Press, Journals Division, 2715 N. Charles Street, Baltimore, MD 21218-4363.

Library Trends is indexed or abstracted in *Current Contents, Library Literature, and PAIS.* The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences–Permanence of Paper for Printed Library Materials. ANSI Z39.48-1984.

Procedures for Proposing and Guest Editing an Issue of *Library Trends*

We encourage our readers to submit ideas for future *Library Trends* themes; issue topics are developed through recommendations from members of the Publications Committee and from reader suggestions. We also encourage readers to volunteer to be issue editors or to suggest others who may be willing to be issue editors.

The style and tone of the journal is formal rather than journalistic or popular. *Library Trends* reviews the literature, summarizes current practice and thinking, and evaluates new directions in library practice. Papers must represent original work. Extensive updates of previously published papers are acceptable, but revisions or adaptations of published work are not sought. Although *Library Trends* is not formally peer-reviewed, guest editors invite articles for submission, which are then critically reviewed by both the guest editor and journal editor. **Unsolicited articles are not accepted**.

An issue editor proposes the theme and scope of a new issue, draws up a list of prospective authors and article topics, and provides short annotations of each article's scope or else gives a statement of philosophy guiding the issue's development. Please send your ideas, inquiries, or prospectus to Library Trends, GSLIS Publications Office, 501 E. Daniel Street, Champaign, IL 61820-6211.

Now Available

INDEXING AND ABSTRACTING IN THEORY AND PRACTICE

3rd edition

By F. W. Lancaster

"Lancaster's work stands by itself as a superior textbook for a course on indexing and abstracting, or as a supplement for any number of courses in related areas (e.g., database management, information retrieval, introductory organization of information, etc.)." —Technicalities

AWARD-WINNING TEXT UPDATED

The third edition of this widely respected manual of best practice is even more comprehensive than the first edition, which won the 1992 ASIS Best Information Science Book Award. Fully revised to address changes since 1998—especially in the areas of multimedia sources, text searching, automatic indexing, and the Internet—this edition is illustrated throughout with useful indexing and abstracting guidelines for student and practitioner. Covered areas include:

- indexing principles and practice
- pre-coordinate indexes
- consistency and quality of indexing
- types and functions of abstracts
- writing an abstract
- evaluation theory and practice
- approaches used in indexing and abstracting services
- indexing enhancement
- natural language in information retrieval
- indexing and abstracting of imaginative works
- databases of images and sound
- automatic indexing and abstracting
- the future of indexing and abstracting services

In addition to use as a text, *Indexing and Abstracting in Theory and Practice* holds value for all individuals and institutions involved in training for information retrieval and related activities, including practicing library and information professionals, database producers, those engaged in portal design, and all professionals involved in knowledge management in general.

ISBN 0-87845-122-6 • 464 pp • cloth • $57.50

Order forms available at: http://puboff.lis.uiuc.edu

Credit orders (Visa, Mastercard, American Express) may be placed by phone: 1-217-333-1359; fax: 217-244-7329; or e-mail: lis-pubs@uiuc.edu

Mail checks payable to the University of Illinois to GSLIS Publications Office, 501 E. Daniel Street, Champaign, IL 61820-6211. *Price does NOT include shipping.

LT-55-1-cover.indd 2 8/14/06 1:54:46 PM

LIBRARY TRENDS

Research Methods

Lynda M. Baker, Issue Editor

Volume 55
Number 1
Summer 2006

Research Methods

Contents

Introduction

Lynda M. Baker

In 1977 Dervin admonished the library and information science (LIS) professionals to stop measuring library activities and start looking at the people who use the library to determine how they use it, how they find information, and how the information helps them. I suggest that this article was a "wake-up" call, challenging our field to adopt new research methods that would allow us to learn more about our clients than about, for example, the number of items circulated. LIS researchers have responded to this call. Powell (1999) and McKechnie and colleagues (2002) have documented the increasing use of research methods adopted from other disciplines. Over the years, *Library Trends* has devoted several issues to research methods. This issue joins the earlier ones and provides information on a variety of traditional and "not so traditional" research methods.

Before describing briefly each contribution, it is important to define "research methods" because as Williamson, Burstein, and McKemmish (2000) pointed out, research methods and data collection techniques are sometimes difficult to distinguish. For example, observation can be both a method and a data collection technique. These authors state "a research method provides a design for undertaking research, which is underpinned by theoretical explanation of its value and use" (p. 11). Data collection techniques are part of the method.

For this volume of *Library Trends,* each author was invited to describe a particular research method and include examples of its use in LIS studies. Articles in this issue are arranged alphabetically by research method and include case study, content analysis, critical incident, discourse analysis, ethnography, evaluation research, life history, longitudinal design, meta-analysis, observation, observation of babies and toddlers, and systematic reviews.

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 1–3
© 2006 The Board of Trustees, University of Illinois

Using a multiple–case studies method, Zach studied the information-seeking behavior of orchestra and museum administrators. Before describing her study, Zach provides a definition of a case study, its historical development, and its use in LIS research. She then briefly describes her study, including sample selection, data collection, and analysis, and her findings.

White and Marsh define content analysis and outline its roots. They describe the procedures involved in both quantitative and qualitative content analysis and provide detailed information on coding and analyzing the data. These authors include two valuable tables: one lists examples of content analysis in LIS research from 1991 to 2005, while the other summarizes the characteristics of qualitative and quantitative content analysis.

Radford has used critical incident technique in her studies of reference encounters. In her article she briefly describes the essence of critical incident technique and illustrates its value in her study of the perceptions of fifth and seventh grade students' encounters with public librarians. A copy of the questionnaire and the instructions to the people who administered it are included in her article.

Budd describes the two major families of discourse analysis, including linguistic-based analysis and culturally or socially based discursive practices. The potential of both families for LIS inquiry and examples from LIS literature are discussed.

Williamson discusses the undertaking of research using a constructivist philosophical framework and ethnographic techniques. A brief discussion of positivism and interpretivism is followed by a section on ethnography. She includes examples from two of her studies: the information-seeking behavior of women with breast cancer, and the information-seeking behavior of online investors.

In his article on evaluation research, Powell outlines reasons for conducting this type of research. After reviewing the general principles and types of evaluation research, he provides information on planning and conducting this type of study, data analysis, and writing the evaluation report. He also includes a list of additional readings on evaluation.

Labaree explores the use of life histories as a research method and the ways it can contribute to new understandings about the experiential relationships between libraries and clients. He covers the essential elements of life history research, describes how to design this type of study, and examines issues related to organization insiderness and internal validity and textual authority.

Davis defines "longitudinal design" as a flexible research approach that can be applied to a wide range of topics involving change over time. She illustrates the use of this type of research in her study of leaders who emerged in the archival profession during the 1980s when archivists developed the

first set of descriptive standards in response to trends in the automation of library cataloging.

In his article on meta-analysis, Saxton provides an explanation of meta-analysis and briefly describes its application in LIS studies. He also provides guidelines for reporting quantitative research, which will enhance the ability of future researchers to perform a meta-analytic study.

While there is considerable literature on observation, the focus of Baker's article is on the roles researchers can adopt in their attempts to gain an in-depth understanding of people in their natural environment. While LIS researchers have played various roles, no instances of complete participation were uncovered in the literature.

McKechnie discusses the practicalities of implementing participant observation in storytime programs for very young children. Included in this article is a list of recommended observation, child development, and research methods texts.

McKibbon outlines the steps involved in conducting a systematic review. This type of review has been widely acclaimed in the health sciences field and is beginning to receive attention by LIS researchers.

As in any edited work, there is some duplication of material. For example, a number of authors have discussed the issues of validity and reliability in qualitative research. No attempt was made to reduce the overlap because, as editor of this volume, I assume that readers will consult a specific article that relates to her/his individual interest. We hope that this volume of *Library Trends* sheds new light on the various research methods described therein. For any LIS researcher, this volume contains a wealth of information not only in the description of each method and its use in LIS studies but also in the numerous citations to seminal works on each research method.

References

Dervin, B. (1977). Useful theory for librarianship: Communication, not information. *Drexel Library Quarterly, 13*(3), 16–32.

McKechnie, L. E. F., Baker, L., Greenwood, M., & Julien, H. (2002). Research methods trends in human information literature. *New Review of Information Behavior Research, 3,* 113–125.

Powell, R. R. (1999). Recent trends in research: A methodological essay. *Library & Information Science Research, 21*(1), 91–119.

Williamson, K. (with Burstein, F., & McKemmish, S.). (2000). Introduction to research in relation to professional practice. In K. Williamson, *Research methods for students and professionals. Information management and systems. Topics in Australasian Library and Information Studies, No. 16* (pp. 5–23). Wagga Wagga, NSW: Centre for Information Studies, Charles Sturt University.

Using a Multiple–Case Studies Design to Investigate the Information-Seeking Behavior of Arts Administrators

Lisl Zach

Abstract

The case study method, and in particular the multiple–case studies design, offers LIS researchers a proven tool for achieving a deep understanding of a specific phenomenon—for example, the information-seeking behavior of a particular user group. Although the case study method has been dismissed by critics who question the rigor of the approach, numerous studies over the past twenty years have demonstrated that the case study method can be used successfully to probe beneath the surface of a situation and to provide a rich context for understanding the phenomena under study. This article summarizes the application of the multiple–case studies design, in which a literal and theoretical replication strategy is used to identify consistent patterns of behavior and to uncover new and/or divergent themes. The motivation behind arts administrators' decisions to seek information is investigated using this approach and examples are given of sample selection, data collection, and analysis. Specific issues associated with the case study method are identified and practical steps used to address them are suggested.

Introduction

Since the early 1980s, when Raya Fidel (1984) published her seminal article on the case study method, case studies have become familiar tools for library and information science (LIS) researchers and have been used successfully to investigate a far-reaching range of topics and users. The case study represents a specific tradition within the qualitative research paradigm (Creswell, 1998) and "attempts, on one hand, to arrive at a comprehensive

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 4–21
© 2006 The Board of Trustees, University of Illinois

understanding of the event under study but at the same time to develop more general theoretical statements about regularities in the observed phenomena" (Fidel, 1984, p. 274). Because case studies are intended to take the reader of the research into the world of the subject(s), case studies can provide a much richer and more vivid picture of the phenomena under study than other, more analytical methods (Marshall & Rossman, 1999).

Like other traditions within the qualitative research paradigm, case studies are used primarily when researchers wish to obtain an in-depth understanding of a relatively small number of individuals, problems, or situations (Patton, 1990). Weick (1979), writing about research in organizations, presented a clear description of the tension among the three primary goals of research: generality, accuracy, and simplicity (by which he meant not only the simplicity of the study but also the understandability of the results). He said that generality is bought at the cost of accuracy—that while a broad study (such as a widely distributed survey) may produce results that can be applied at a general level to a large number of organizations, the results are unlikely to present an accurate description of any one organization. This tension exists in case study research as well—depth of understanding about the phenomena under study is bought at the cost of "confident generalizations" (Patton, 1990, p. 53) about the applicability of the results to individuals, problems, or situations outside of the study parameters.

This article seeks to describe the nature of case study research, specifically the use of the multiple–case studies design described by Yin (1994), and to give an example of its application in a study of the information-seeking behavior of senior arts administrators.

What Is a Case Study?

A case study is "an exploration of a 'bounded system' . . . a program, an event, an activity, or individuals" (Creswell, 1998, p. 61). The concept of a case study comes from the practice of law, in which the unit of analysis is a single case before a court. We are familiar with the use of case studies as pedagogical tools in law and business (for example, the Harvard Business School case study approach). Sigmund Freud made the case study famous as a method of documenting his observations of patients in psychoanalysis (Breuer & Freud, 1895). Often a case study recounts a rare or unusual condition or event, but it may also be a description of a classic situation that can be used as a model or exemplar.

Historical Development

The case study method as practiced in LIS research today has its roots in the social sciences, especially in sociology. In 1992 *Current Sociology,* the journal of the International Sociological Association, published an issue devoted to the development and use of the case method in sociology (Hamel, 1992). Any student of the case study method would be well served to review

its history, and the annotated bibliography contained in that issue provides an excellent source of further readings (Dufour & Fortin, 1992).

In the early part of the twentieth century, case studies were referred to as tools in the realm of social work; by the 1930s the case study method was accepted as a procedural alternative to the statistical method among researchers at the University of Chicago (Platt, 1992). Case studies were seen as valuable because of the rich context in which they placed the subjects of the inquiry. Unlike statistical studies, case studies were perceived to allow the researcher to see beneath the surface of the situation into personal meaning (Burgess, 1928). However, proponents of the statistical method gained momentum, and by the middle of the century case studies were largely relegated to the role of preliminary or exploratory research, where they were used to "suggest hypotheses for more systematic investigation" (Platt, 1992, p. 28).

In the 1960s a new generation of researchers became interested in qualitative methods, especially as an approach for developing theory (Glaser & Strauss, 1967). These researchers continued to face critics who raised issues concerning the reliability (the extent to which repeating the same procedures under the same conditions would produce the same results) and validity (the extent to which the research matched its stated goals) of the findings of their studies. Case studies in particular were criticized because of the lack of rigor of the research methods employed and the degree to which personal bias, either of the participants or of the researcher, could influence the findings and conclusions. The way in which case studies were being carried out led Simon, in a textbook on basic research methods, to conclude that "the investigator makes up his procedure as he goes along" (1969, p. 276).

Another reason that case studies were particularly vulnerable to criticisms of this nature was the use of participant observation as a method of data collection instead of the more accepted approach of structured interviews or questionnaires. In any study that relies on observed behavior, there is always the possibility that the very act of studying the behavior will alter it.[1] With participant observation, not only does the researcher record the behavior, he also may play a variety of roles in the activities being studied. The advantage of this approach is that the participant-observer may gain access to groups or situations otherwise closed to researchers; he can also be opportunistic about following new research directions as they present themselves. The obvious drawback to the approach is the potential for bias, both in data collection and analysis.

As the interest in qualitative methods revived, researchers created a new language to describe certain concepts related to reliability and validity and to address the concerns over the lack of a rigorous research structure. Guba (1981) proposed "trustworthiness" as a surrogate measure for validity and reliability in naturalistic inquiries. "Trustworthiness" in this context is a belief system that informs the whole way in which the researcher

approaches a research study (Guba & Lincoln, 1982). By structuring the study to address the four aspects of trustworthiness—that is, truth value, applicability, consistency, and neutrality—the researcher hopes to achieve the following outcomes:

- Credibility: the credibility of any qualitative research study speaks to the issue of whether the findings are plausible; this in turn rests on the steps taken during the whole process of data collection and analysis. Key among the factors that ensure credibility are the completeness of the data collection, the use of multiple analytical perspectives, and member checks to confirm the accuracy of the conclusions drawn (Yin, 1994).
- Transferability: the transferability of a research study addresses the question of whether the findings are "context-relevant" or subject to noncomparability because of situational uniqueness (Guba, 1981, p. 86). To provide a context for evaluating the transferability of the findings, the researcher should use theoretical and/or purposive sampling and develop a thick description of the data that can be reviewed by others.
- Dependability: the goal of confirming the dependability of the data to ensure the stability of the findings is a challenging one for researchers; overlapping methods of data collection and/or stepwise replication are the recommended approaches (Guba, 1981). However, due to practical limitations, many researchers must rely primarily on establishing a good "audit trail" of project documentation that can be followed by others.
- Confirmability: to avoid the effects of investigator bias, steps should be taken to collect data from a variety of sources and, if possible, by researchers with different perspectives. When these steps are not possible, the researcher should rely on "practicing reflexivity," which Guba describes as revealing the researcher's own assumptions to his audience (Guba, 1981, p. 87). This can be done by documenting personal reactions and beliefs about the data.

By specifically addressing the concerns of critics, researchers working within the qualitative research paradigm hoped to gain acceptance for their methods. However, because the case study approach typically involves "intense analyses of a small number of subjects rather than gathering data from a large sample or population" (Powell, 1997, p. 49), a further concern of quantitative researchers was the lack of generalizability of the results. It is interesting to note that one of the areas in which case study methods became and have remained popular is in the area of organizational research, where the focus is on understanding a particular work environment or structure and not necessarily on predicting results in other areas (Van Maanen, 1988). For those who have wanted to make generalizations based on case studies, some researchers have attempted to develop methods for quantifying data from case studies, but most agree that it is difficult if not impossible to generalize from case studies to a wider population.

Use in LIS Research

In 1984 Fidel published her article describing "the first time that the case study method has been used in library and information research to differentiate broad patterns of behavior" (p. 273). In it she defined the case study as a specific type of field study. As such, she explained, researchers using this method would be influenced in their data collection by what they found in the field. Data collection would be accomplished using approaches determined by the subject matter; these could include direct observations, interviews, or document analysis. An iterative approach to data collection and analysis was recommended so that the results of previous analysis could direct further investigation. The desired outcome of the investigation was to be both "comprehensive understanding" and the development of "general theoretical statements about regularities" (p. 274). In her description of the method she used, Fidel addressed both the issues of reliability and validity and acknowledged that they cannot be ensured in the case study method, but she asserted that other methods exist to ensure the rigor of the approach. She also addressed the issues of access to subjects, study effect, participant bias, and observer bias.

With this article providing the heretofore missing guidance needed by LIS researchers to apply the case study method, use of this approach grew dramatically.[2] Case studies have been used in LIS research to investigate groups of library users and nonusers as diverse as children, college students and faculty, professionals (doctors, lawyers, managers, etc.), the culturally disadvantaged, and persons in hospitals and correctional institutions. The method has also been used to study libraries as institutions. "Indeed," wrote Busha and Harter, "the case [study] approach is particularly applicable in inquiries concerned with the role of libraries as social institutions—that is, their social control, performance, and impact on society in general and special groups in particular" (1980, p. 152).

A recent study by Donald Case on survey methods used in research on information seeking, needs, and behavior describes the "simplicity and groundedness" of the case study method, comparing it to "more elaborate methods" (2002, p. 178). Case focuses on the case study as an approach primarily used to delve deeply into a single subject, as in the example he provides of research by Carol Kuhlthau—a longitudinal study following a single securities analyst through his on-the-job learning process (Kuhlthau, 1999). During each of several stages in the analyst's career, Kuhlthau conducted in-depth interviews with him and used the results to draw conclusions about the role experience plays in information-seeking behavior. This type of extended contact with the subject is extremely rare in LIS case study research and provides a model of how the method can be used to explore the rich context of the phenomena under study.

Kuhlthau (1999), when describing the value of the method, reverses the more traditional concept of using a case study as an exploratory approach

to identify characteristics that may lead to further research questions. Rather, she suggests that using the case study is a way "to gain insight into some of the questions raised in prior, more quantitative, studies" (p. 411). In order to provide a thorough understanding of the phenomena under study, she advocates the use of a mixed-method approach; both qualitative and quantitative research methods can be used to compliment each other (Kuhlthau, 1993). Other approaches that can be used to develop the rich context of a study include methodological and theory triangulation (Patton, 1990), that is, the use of multiple sources of data or evidence, for example, observations, interviews, documents, and even surveys (Solomon, 1997), and multiple analytical perspectives, for example, different cultural or theoretical views (Yin, 1994).

Multiple–Case Studies Design

While much case study research focuses on a single case, often chosen because of its unique characteristics, the multiple–case studies design allows the researcher to explore the phenomena under study through the use of a replication strategy. Yin (1994) compares the use of the replication strategy to conducting a number of separate experiments on related topics. Replication is carried out in two stages—a *literal replication* stage, in which cases are selected (as far as possible) to obtain similar results, and a *theoretical replication* stage, in which cases are selected to explore and confirm or disprove the patterns identified in the initial cases. According to this model, if all or most of the cases provide similar results, there can be substantial support for the development of a preliminary theory that describes the phenomena (Eisenhardt, 1989).

In the multiple–case studies design, there are no hard-and-fast rules about how many cases are required to satisfy the requirements of the replication strategy—Yin suggests that six to ten cases, if the results turn out as predicted, are sufficient to "provide compelling support for the initial set of propositions" (1994, p. 46). Yin goes on to say that, since the multiple–case studies approach does not rely on the type of representative sampling logic used in survey research, "the typical criteria regarding sample size are irrelevant" (p. 50). Instead, sample size is determined by the number of cases required to reach saturation, that is, data collection until no significant new findings are revealed. The sample participants should be selected explicitly to encompass instances in which the phenomena under study are likely to be found. This approach to sample design is consistent with the strategy of homogeneous sampling, in which the desired outcome is the description of some particular subgroup in depth (Patton, 1990).

Application of Multiple–Case Studies Design

The following sections provide an example of the application of a multiple–case studies design to investigate the information-seeking

behavior of arts administrators. The study addressed the following research questions:

- How do arts administrators go about getting the information they want?
- How do they determine that they have "enough" information?
- How much effort are they willing to invest in seeking information?

Sample Selection

For the study, a sample pool of arts administrators was drawn from two of the disciplines within the arts field: symphony orchestras and art museums.[3] These two disciplines were chosen because they represent different traditions in arts administration and attract administrators with different educational and professional backgrounds. These differences provided the opportunity for both the literal and the theoretical replication process. The final sample group included seven orchestra administrators and five museum administrators. The sample comprised experienced practitioners in their fields: the average number of years in the field was twenty-eight. What little research that has been done on arts administrators as a group shows that they are notably well educated (DiMaggio, 1988). This conclusion was confirmed in this study: all but one administrator have at least one advanced degree—two have Ph.D. degrees.

Access to the sample group was gained through personal contacts. All the administrators contacted expressed an initial willingness to participate in the study, although several later withdrew because of scheduling constraints. Orchestra administrators were contacted first because of the researcher's prior work relationship with these individuals. Orchestra administrators were selected (as far as possible) to fulfill the literal replication phase of the multiple–case studies design; the museum administrators were selected to explore and confirm or disprove the patterns identified in the initial interviews (theoretical replication). Museum administrators were identified by the orchestra administrators or through the researcher's personal contacts. Ultimately, the specific participants were selected based on their availability at the time of data collection. This approach is consistent with the concept of open sampling, in which the selection of specific interviewees or observational sites within a target group can be indiscriminate since the purpose is to collect as much data as possible to guide the early phases of theory development (Strauss & Corbin, 1998).

Data Collection

For the purpose of this study, a "case" was defined as a single, in-depth interview with an arts administrator. Data were collected from the twelve arts administrators over a four-month period using a pre-tested interview protocol that included twenty-five questions focusing on specific information-seeking tasks, information sources, stopping criteria, and general

information-seeking style. Each question was mapped to one or more of the main research questions. After the interview with the first orchestra administrator was conducted, the results were transcribed and analyzed before the next group of interviews was scheduled. The next three interviews with orchestra administrators were conducted within a one-week period; one of these interviews provided results quite different from the other two. The fifth interview conducted with an orchestra administrator, however, was consistent with the aims of *literal replication,* that is, the interview substantially confirmed information collected in three of the four earlier interviews. The next group of three interviews, which were with museum directors, was used both to investigate any museum-specific behavior patterns and to confirm or disprove the patterns of behavior identified in earlier interviews (*theoretical replication*). The remaining four interviews (two in each user group) were used to explore and/or contrast the patterns identified in the earlier interviews. The final four interviews did not produce any new concepts; they did provide the opportunity to explore specific concepts in more depth and to deepen the understanding of the phenomena. After completion of the fifth interview with museum administrators, it was determined that no new information had been obtained. Data collection was therefore discontinued. A diagram of the process used in the study of arts administrators is shown in Figure 1.

During the course of the first five interviews with arts administrators, several slight revisions were made to the interview protocol to adjust or reorder questions that seemed confusing or unproductive. After completing the preliminary analysis of the first set of interviews, four new questions were added to the protocol that allowed the researcher to explore certain new concepts during the theoretical replication phase. These questions focused on the reasons why administrators make the decision to look for information in the first place and how they choose specific individuals as sources of information. Although these issues had been addressed to some degree in the initial protocol, it became obvious during the literal replication phase that these sections of protocol needed to be expanded.

The ability to adjust the data collection as a result of insights obtained during the early phases of the research process highlights one of the key advantages of the multiple–case studies design. The research questions used to guide this research concentrated on the "how" of the arts administrators' information-seeking behavior; after the initial set of interviews were completed and used to define the norm, data collection during the theoretical replication phase could focus on the "why" of their individual behaviors. This allowed the researcher to explore the reasons for these perceived differences and to build explanations for them based on responses to an expanded set of questions.

The multiple–case studies design provides a rigorous approach for collecting and analyzing data. As shown earlier in Figure 1, the replication

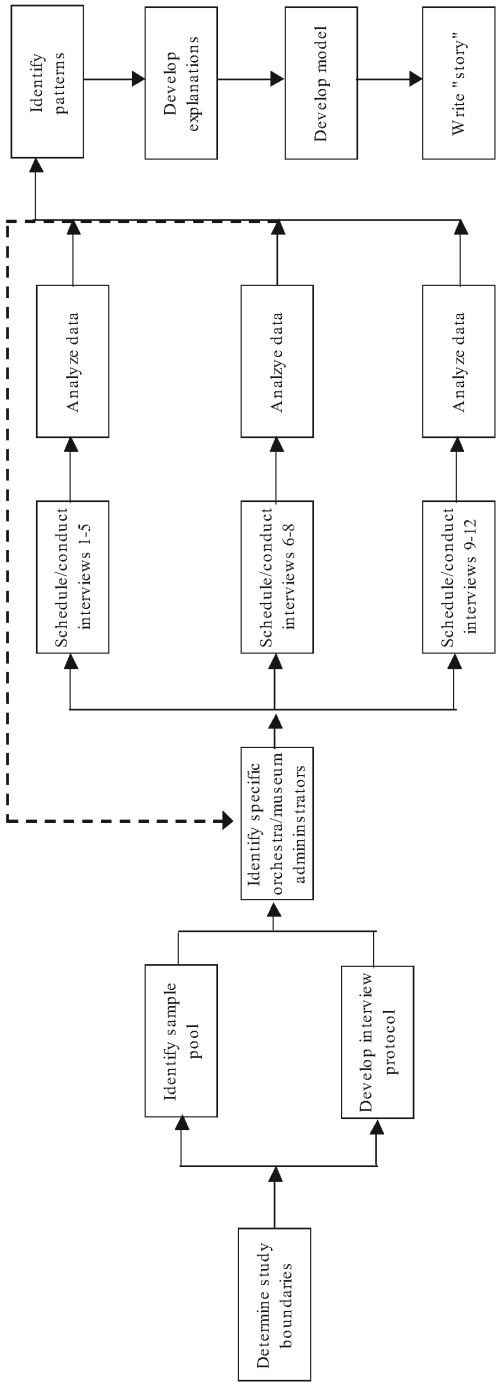

Figure 1. The Multiple-Case Studies Design.

Adapted from Yin, R. K. (1994) *Case Study Research.*

strategy allows the researcher to identify possible patterns in the data and explore them by returning to the field for more data. Conscientious application of these techniques ensures that explanations for the phenomena under study developed from the data are verified during the course of the research process. This iterative process of data collection, analysis, comparison, and revision during the entire study is referred to as the "constant-comparative" method (Strauss & Corbin, 1998).

Data Analysis

The main approach to data analysis involved a detailed analysis of the interview transcripts. As the first step in this sequential process, notes from each interview, made both during the interview and immediately after it, were reviewed; highlights or new concepts were identified. Next, the transcript from each interview was reviewed and coded.[4] As the process continued, each new interview was compared to the previous ones for confirming or disconfirming evidence; earlier interviews were reanalyzed in the light of new concepts identified in later interviews. Because the multiple–case studies design encourages the researcher to analyze the data from earlier interviews before scheduling and conducting the later ones, the analytic process itself influences the emphasis placed on certain questions during the later part of the process.

Preliminary patterns describing the factors that influence information-seeking behavior were developed based on concepts identified during the literature review. These patterns were augmented by concepts that came out of the first group of interviews. The first one or two interviews from each of the two user groups were an especially rich source of new concepts. For example, during one of the early interviews, an administrator expressed a completely unexpected opinion. In response to this administrator's strongly held position, the researcher included a new question in the interview protocol on the use of the organizational mission in the information-seeking process. During the theoretical replication phase, the researcher found confirming evidence for the phenomenon of mission alignment as a factor influencing the decision to seek information, although only two other administrators exhibited the same extreme position.

As the interview process continued, predictable patterns began to emerge, allowing the researcher to form an early interpretation of the nature of the information-seeking process used by senior arts administrators. To the extent that the patterns found in the data from each additional interview matched the early interpretation of the process that had been developed, the internal validity of that interpretation was strengthened.

During the pattern-matching process, the data collected were organized to support plausible explanations about the nature of information-seeking and stopping behavior among senior arts administrators. Based on these explanatory patterns, the initial model describing the information-seeking

process was developed. During the explanation-building process, some individual situations were identified that did not appear to fit the model being developed. For these cases, it was necessary to review the data and discover what intervening conditions, if any, might exist that could explain specific differences in behavior.

Throughout the analytic process, multiple perspectives were used to interpret the data and to provide theory triangulation. Specifically, the data were reviewed from traditional management and arts management perspectives as well as from the perspective of an LIS researcher. The use of these perspectives helped to explain otherwise anomalous behavior on the part of individual arts administrators and to reduce the risk that any single interpretation of the data would shape the results. Data triangulation was obtained by the fact that the informants themselves came from two separate user groups and represented different types of organizations. As an ongoing check in the process, each interview was reviewed specifically to look for evidence that ran contrary to the norm; no disconfirming evidence was found that could not be explained by specific intervening conditions. Finally, member checks with each of the study participants were used to confirm the conclusions of the study and to guard against the possibility of researcher bias and reactivity.

Findings

Data supporting the findings of a case study may be presented in a number of ways, including making a matrix of categories and placing evidence within such categories, creating arrays—flowcharts and other devices—for examining the data, or tabulating the frequency of different events. However, one of the most powerful tools that the writer of a case study report can use is the evidence of the participants' own words to "tell the story." This brings the reader into the participants' world and provides a rich context for understanding the phenomena under study.

During the interviews with the arts administrators and subsequent analysis of the data, three major themes emerged that had not been identified in the original research questions. Of these, one in particular—the motivation for seeking information—was the result of evidence uncovered during the literal replication phase that did not fit into the expected framework. This concept was then explored in depth during the theoretical replication phase, and the results are described below.[5]

The Decision to Seek Information Based on the early interviews, it became apparent that arts administrators do not assume a priori that their tasks or decisions will require a formal information-seeking process. While admitting that some situations may involve formal information seeking, those situations appeared to be the exceptions rather than the rule. Administrators rely heavily on their personal experience, previous knowledge, and

randomly acquired information in order to perform their jobs. Some of them question whether a formal process of information seeking provides them with valuable or even useful results. During the theoretical replication phase, therefore, the researcher focused on an issue not addressed in the initial foreshadowing questions: What motivates an administrator to engage "purposefully" in the search for information? (Marchionini, 1995).

None of the administrators had formal guidelines for when he would look for information or when he would not. One administrator said: "I intuitively know when I can make a decision just on gut, because I know enough information already, and when to seek out additional information. But it's really, really hard to write a manual of how to do that. I think that is something that may separate successful executives from unsuccessful [ones]."

Based on data collected and analyzed during the literal replication phase, a number of specific situations were identified for which administrators generally agreed that they might use a formal process to look for information. These included budgeting and other financially driven activities, long-range planning, and audience or market research projects. There was also general agreement that administrators would be more likely to use a formal process for questions or situations that were either new to the organization or outside their personal areas of expertise: "I guess the other situation is where you're trying to make a decision, and it's not only a new experience for us, but you either can't find anybody else with credibility from whom to get information, or maybe it really is a new issue." Another administrator added the caveat that looking for information does not necessarily mean that one will use it: "We go through the motions of gathering information about things that we're closed-minded about, and where we think we already know what's best. We'll still get it, but we'll either do one of two things. We'll ignore it, or we'll say, well you have to consider the source and factor that out or filter it."

Mission-Driven Information Seeking Alignment to mission emerged during the literal replication phase as a factor influencing the decision to seek information for one administrator. This administrator maintained that his need for information was both defined and limited by asking the question, "How does this task/decision relate to my organization's mission?" If it did, then he needed relatively little information on which to make his decision; if it did not, then he did not even consider the decision, regardless of other circumstances surrounding it.

> The experience here is that the overall criteria for institutional success is so firmly built into my mindset, and everyone's mindset, that it actually . . . helps force the right decision, the right questions get asked. . . . By framing the objective of the project as we have, . . . it immediately provides a framework for thousands of mini-decisions . . . and

> understand it framed the criteria in advance. . . . I'm never conscious about what do I need to know, what information do I need? I'm very comfortable, I'm extremely comfortable with my instinct, and I hardly ever get bogged down in information.

One of the two other administrators who supported this position illustrated his point about the use of the mission to influence the need for information by explaining that all he had to do to convince his board to approve the loan of artwork to a nontraditional exhibit was to assure them that this exhibit would be completely consistent with the organization's mission.

Five administrators agreed that, although they referred to the organization's mission statement as a guideline in decision making, it was not an absolute determinant of behavior nor a factor strongly influencing information seeking. One other administrator, who described himself as completely mission-driven, summed up the value of using a mission statement to focus information seeking: "I can just provide a kind of . . . rule of thumb, and I think it's true. If you're mission-driven, you're going to come out better in the end—at the end of the day. So if a symptom of being non-mission-driven, or less-mission-driven, is looking for more information, then if you're to flip it around and say the people who are looking for lots of information probably don't have their ducks lined up."

Information Seeking as Consensus Building The second new theme to emerge as a factor influencing an administrator's decision to seek information was the use of information seeking as a means for consensus building. Several administrators made comments such as, "I knew where I wanted to go with this, but. . . ." and then went on to describe how they had set up an elaborate information-seeking process primarily or even solely to involve various constituent groups and to develop a sense of buy-in:

> Sometimes you know the answer before you start. You're just building the case for it. . . . And as I'm thinking further about this, because we want consensus, so we think of what information we want, we get some information back, we might redefine the information, and then what happens is if we as a small group decide that we believe we have the answer, then we need to look at the information again, because we want to build consensus and then how do we present the information so that it's clear and understandable and honest, and you get a presentation you can make to someone else to help build consensus.

Seven of the twelve administrators gave some example of information seeking being used as a tool for bringing people together around an issue or for moving a decision in a desired direction. However, none of the three administrators who considered themselves "mission-driven" mentioned this use of information seeking; rather, they described a much more focused approach to decision making, often centering on a small group of senior staff who depended very little on outside input.

Although the interview protocol did not specifically address the administrators' roles in providing information to others, four administrators interviewed during the theoretical replication phase volunteered information about how they see themselves in those roles when asked about their decision to seek information. Not surprisingly, the administrators who believed most strongly in consensus building also believed in disseminating or sharing information with others in their organizations, even if they questioned how this effort was received. "And consensus for me is critical. I am not an autocratic leader. . . . For me, communication, information sharing, is critical for being able to move the organization forward. . . . I think the other part is that I believe that this [information seeking] is multidirectional. I seek a lot of information. I share a lot of information."

Although the concept of information seeking as a means for consensus building was mentioned early in the interview process, the possible relationship between level of information seeking and specific organizational cultures, that is, mission-driven and consensus-based, was not identified until late in the analysis phase. This relationship between organizational culture and executive behavior has been explored by many researchers (Mintzberg, 1973; Dees, 1998; Martin, 2002), but the relationship between specific nonprofit cultures and information seeking is an area for future research. From the indications provided by this study, it would appear that these two types of organizational cultures exert opposite influences on information-seeking behavior.

Considerations and Concerns

In her 1984 article Fidel identified several potential problems associated with the case study method. These include access to participants, study effect, participant bias, and observer bias. These issues remain as challenging now as they were twenty years ago, and the practical steps used to address them in the study of arts administrators are described below.

Access to Subjects

As Fidel and others have repeatedly suggested, one of the challenges of any case study research is "getting in" (1984, p. 285). Various authors have offered suggestions about access to participants, but in the end, studies of many different populations and environments have shown that the most effective approach is through personal contacts. Because of the researcher's previous familiarity with the orchestra field, administrators from this discipline were selected first and used to populate the *literal replication* phase of the design. Since the researcher knew all the orchestra administrators in this sample personally, this contact was relatively straightforward.

Since the researcher did not have the same level of access to museum administrators as to orchestra administrators, an entry strategy had to be developed and appropriate contacts needed to be identified and asked to

help identify interview candidates. Two sources of referrals were used to gain access to museum administrators: (1) the orchestra administrators already being interviewed for the study, and (2) personal contacts known to the researcher who serve as museum board members. The preferred approach was to ask each orchestra administrator being interviewed to recommend a colleague in a museum who might be willing to participate in the study. If a museum administrator contacted through an orchestra administrator could not participate, then the second, and usually more attenuated referral source, was used to recruit participants.

In all cases, the personal credibility of the researcher and/or the person making the referral was essential to secure access. Although it is certainly possible to obtain access to many user groups without previous personal knowledge, it is often challenging. In the absence of familiarity, the affiliation with an institution, such as a university, often provides the necessary credibility for the researcher. Also, without previous personal knowledge, it may take longer for participants to "open up" to the researcher and share candid opinions in response to questions.

Study Effect

As discussed above, one of the criticisms of the case study method has long been that the very act of studying a phenomenon may alter it. In the study of arts administrators, two of the participants specifically mentioned that the very nature of the questions caused them to think differently about their information-seeking behavior than they had before. Because arts administrators do not consider information seeking to be an important part of their decision-making process, they do not think about it as a conscious activity. When asked to do so, they begin to construct reasons for their behavior that may not be accurate.

Because the researcher was not in a position to observe actual information-seeking behavior over time, it was necessary to take the arts administrators' descriptions of their processes at face value. Ideally, the researcher would seek an external evaluation of the accuracy of the descriptions. However, direct observation of the information-seeking process, although desirable, was not realistic, since the process may take place over an indefinite period and is often co-mingled with other tasks; it is also too intrusive an approach when studying the behavior of individual administrators. Because of the nature of the group being studied, the use of secondary sources, diaries, and/or activity sampling was also not appropriate, nor could a questionnaire be designed that would provide reliable results.

Since it is impossible to eliminate completely the risk of the study effect on the participants, extra care must be taken during the data collection and analysis process to ensure that any unusual behavior is identified and evaluated to determine whether it is caused by an outside influence. In the absence of multiple sources of evidence about the behavior of individual ad-

ministrators, the researcher relied on a limited amount of data and theory triangulation to reduce the risk of misinterpreting the evidence or placing undue importance on anomalous data.

Participant Bias

Participant bias presented a very small problem in the study of arts administrators. In general, the arts administrators interviewed for the study were easy to talk to and enjoyed describing their own views and experiences. Two administrators directly questioned the premise of the study but admitted to finding the questions interesting and thought-provoking. At the other end of the spectrum, some administrators said things like, "This is fun" and "This is just what I need." Since the study was interested in the opinions and perceptions of the administrators, the fact that they were predisposed to dismiss information seeking as an important activity only provided additional material for analysis.

Observer Bias

The obvious downside of personal knowledge of a particular participant group is the potential for bias when dealing with it. On the other hand, this knowledge can also provide theoretical sensitivity. Strauss and Corbin define sensitivity as "having insight into, and being able to give meaning to, the events and happenings in data" (1998, p. 46). Sensitivity is a quality that helps a researcher to recognize what may be significant in the data and/or to identify inconsistencies between an individual's behavior and standard practice. Theoretical sensitivity, according to these authors, may be derived from the relevant literature, professional and/or personal experience, and the analytical process itself. In this case, the initial theoretical sensitivity was brought to the situation through both the relevant literature and the researcher's own professional experience and personal interests. Having worked as a member of the senior management team in two arts organizations, this researcher has operational knowledge concerning how and where senior administrators are likely to look for the information they want. In addition, the researcher has been involved in advising arts administrators on how to identify and satisfy their information needs. However, no explanation arising from previous experience was included unless it was verified by actual data collected from the field.

Conclusions

The case study method, and in particular the multi–case studies design, provides LIS researchers with a proven tool for achieving a deep understanding of a specific phenomenon—for example, the information-seeking behavior of a particular user group. The strength of the multiple–case studies design lies not only in its ability to demonstrate consistent patterns of behavior but also, and perhaps more importantly, in its ability to uncover new and/or divergent themes. These emerging themes can be explored

through the replication process. This process allows the researcher to probe beneath the surface of the situation and to focus on the "why" of individual behaviors.

Although case studies do present problems to the researcher in terms of access, study effect, and potential sources of bias, these issues can all be addressed by the application of rigorous data collection and analysis techniques. As has been demonstrated by numerous studies over the past twenty years, the case study method can be used not only for exploratory research but also for theory development. The case study method and the rich context that it offers often provide the reader of the research with a much more vivid experience than do other, more analytical methods.

Notes

1. This phenomenon is referred to as the Hawthorne Effect, named after a study of factory workers at Western Electric's Hawthorne Plant in Illinois from 1927 to 1933. The study showed that productivity increased as a result of the very act of studying it, regardless of any changes made in working conditions.
2. The use of case studies in library/information science dissertations almost tripled (up 284 percent) between 1975–79 and 1990–94 (Blake, 2003).
3. For a complete description of the sample selection and other methods used in this study, see Zach (2002).
4. Data were coded using version 1.3.146 of NUD*IST (NVivo) developed by Qualitative Solutions and Research. At the end of the process, 1,753 passages had been coded using 389 terms.
5. For a more complete discussion of information seeking by this user group, see Zach (2005).

References

Blake, V. L. P. (2003). Research methods in LIS dissertations. In M. A. Drake (Ed.), *Encyclopedia of library and information science* (2nd ed., pp. 2513–2523). New York: Marcel Dekker.

Breuer, J., & Freud, S. (1895). *Studien uber hysterie.* Leipzig: F. Deuticke.

Burgess, E. W. (1928). What social case studies records should contain to be useful for sociological interpretation. *Social Forces, 6*(4), 524–532.

Busha, C. H., & Harter, S. P. (1980). *Research methods in librarianship: Techniques and interpretations.* San Diego, CA: Academic Press.

Case, D. O. (2002). *Looking for information: A survey of research on information seeking, needs, and behavior.* New York: Academic Press.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions.* Thousand Oaks, CA: Sage.

Dees, J. G. (1998). Enterprising nonprofits. *Harvard Business Review, 76*(1), 55–67.

DiMaggio, P. (1988). *Managers of the arts: Careers and opinions of senior arts administrators of U.S. art museums, symphony orchestras, resident theatres, and local arts agencies.* Washington, DC: Seven Locks Press.

Dufour, S., & Fortin, D. (1992). Annotated bibliography of case study method. *Current Sociology, 40*(1), 166–181.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review, 14*(4), 532–550.

Fidel, R. (1984). The case study method: A case study. *Library & Information Science Research, 6*(3), 273–288.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication & Technology Journal, 29*(2), 75–91.

Guba, E. G., & Lincoln, Y. S. (1982). Epistemological and methodological bases of naturalistic inquiry. *Educational Communication and Technology Journal, 30*(4), 233–252.
Hamel, J. (Ed.) (1992). *Current Sociology* (Vol. 40). Thousand Oaks, CA: Sage.
Kuhlthau, C. C. (1993). *Seeking meaning: A process approach to library and information services.* Norwood, NJ: Ablex.
Kuhlthau, C. C. (1999). The role of experience in the information search process of an early career information worker: Perceptions of uncertainty, complexity, construction, and sources. *Journal of the American Society for Information Science, 50*(5), 399–412.
Marchionini, G. (1995). *Information seeking in electronic environments.* New York: Cambridge University Press.
Marshall, C., & Rossman, G. B. (1999). *Designing qualitative research* (3rd ed.). Thousand Oaks, CA: Sage.
Martin, J. (2002). *Organizational culture: Mapping the terrain.* Thousand Oaks, CA: Sage.
Mintzberg, H. (1973). *The nature of managerial work.* New York: Harper and Row.
Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.
Platt, J. (1992). "Case study" in American methodological thought. *Current Sociology, 40*(1), 17–47.
Powell, R. R. (1997). *Basic research methods for librarians* (3rd ed.). Greenwich, CT: Ablex.
Simon, J. L. (1969). *Basic research methods in social science.* New York: Random House.
Solomon, P. (1997). Discovering information behavior in sense making: I. Time and timing; II. The social; III. The person. *Journal of the American Society for Information Science, 48*(12), 1097–1138.
Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.
Van Maanen, J. (1988). *Tales from the field: On writing ethnography.* Chicago: University of Chicago Press.
Weick, K. E. (1979). *The social psychology of organizing* (2nd ed.). San Francisco: Jossey-Bass.
Yin, R. K. (1994). *Case study research: Design and methods.* Thousand Oaks, CA: Sage.
Zach, L. (2002). *When is "enough" enough? An investigation of the information-seeking and stopping behavior of senior arts administrators.* Unpublished doctoral dissertation, University of Maryland, College Park.
Zach, L (2005). When is "enough" enough? Modeling the information-seeking and stopping behavior of senior arts administrators. *Journal of the American Society for Information Science and Technology, 56*(1), 23–35.

Lisl Zach is an assistant professor at LSU's School of Library and Information Science, where she teaches in the areas of knowledge management and organization of information. Dr. Zach has spent twenty years working in various areas of administration, as well as conducting applied research to identify ways of satisfying the information needs of user groups as diverse as field artillery officers, nuclear power plant operators, and symphony orchestra managers, staff, and volunteers. Dr. Zach's academic interests include studying the information-seeking behavior of decision makers and investigating ways of measuring and communicating the value of information to organizations. She is also concerned with the methodological issues related to qualitative research, including various approaches to data collection, analysis, and presentation. Dr. Zach holds a Ph.D. in Information Studies from the University of Maryland–College Park, an MBA from New York University, and an MSLS from the University of North Carolina–Chapel Hill.

Content Analysis: A Flexible Methodology

Marilyn Domas White and Emily E. Marsh

Abstract

Content analysis is a highly flexible research method that has been widely used in library and information science (LIS) studies with varying research goals and objectives. The research method is applied in qualitative, quantitative, and sometimes mixed modes of research frameworks and employs a wide range of analytical techniques to generate findings and put them into context. This article characterizes content analysis as a systematic, rigorous approach to analyzing documents obtained or generated in the course of research. It briefly describes the steps involved in content analysis, differentiates between quantitative and qualitative content analysis, and shows that content analysis serves the purposes of both quantitative research and qualitative research. The authors draw on selected LIS studies that have used content analysis to illustrate the concepts addressed in the article. The article also serves as a gateway to methodological books and articles that provide more detail about aspects of content analysis discussed only briefly in the article.

Introduction

As a research methodology, content analysis has its roots in the study of mass communications in the 1950s.[1] Based on a basic communications model of sender / message / receiver, initially researchers emphasized mak-

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 22–45
© 2006 The Board of Trustees, University of Illinois

ing inferences based on quantified analysis of recurring, easily identifiable aspects of text content, sometimes referred to as *manifest content*. Since then, researchers in many fields, including anthropology, library and information studies (LIS), management, political science, psychology, and sociology, have used content analysis. In the process, they have adapted content analysis to suit the unique needs of their research questions and strategies and have developed a cluster of techniques and approaches for analyzing text grouped under the broad term of *textual analysis*. A significant change has been a broadening of text aspects to include syntactic, syntagmatic, and pragmatic aspects of text, although not always within the same study. Merten (as cited by Titscher, Meyer, Wodak, & Vetter, 2000) notes that "the range of procedures in content analysis is enormous, in terms of both analytical goals and the means or processes developed to pursue them" (p. 55). The variants include, for example, besides content analysis, conversational analysis, discourse analysis, ethnographic analysis, functional pragmatics, rhetorical analysis, and narrative semiotics.[2] Although these approaches are alike in their reliance on communicative material as the raw material for analysis, they vary in the kinds of questions they address and in their methods.

This article focuses only on content analysis, not on all forms of textual analysis. It distinguishes, however, between quantitative and qualitative approaches to content analysis since both are used in information studies. Content analysis is a flexible research method that can be applied to many problems in information studies, either as a method by itself or in conjunction with other methods. Table 1 provides a selective list of research studies in LIS using content analysis published within the past fifteen years (1991–2005).

After defining content analysis, the article goes through the basic steps in a content analysis study. It does this first for quantitative content analysis, then notes the variations that exist for qualitative content analysis. Throughout the article draws on the LIS studies in Table 1 for examples. Although only certain aspects of the LIS studies are mentioned in the article, they constitute a rich trove showing the broad applicability of content analysis to many topics. The article closes with a brief bibliographical note leading to sources providing more detail about the content analysis aspects treated only briefly here.

Definition

Not surprisingly, multiple, nuanced definitions of content analysis exist that reflect its historical development. This article accepts a broad-based definition in a recent content analysis textbook by Krippendorff (2004).[3] For the purpose of this article, content analysis is "a research technique for making replicable and valid inferences from texts (or other meaningful

Table 1. Selected Examples of Content Analysis in LIS Research, 1991–2005

Article[a]	Purpose	Data	Type of Analysis[b]	External Coding Scheme
Agosto & Hughes-Hassell, 2005	To describe the everyday life information-seeking patterns of urban young adults	Written activity logs and transcribed semi-structured group interviews[c]	Qualitative	
Buchwald, 2000	To explore the role of a public interest group in federal information policy making	Recorded and transcribed informant interviews, observation notes, group documents	Qualitative	
Croneis & Henderson, 2002	To identify the changing nature of work in academic libraries, focusing on positions emphasizing electronic and digital resources	Job advertisements in *College & Research Libraries* for a ten-year period	Qualitative	
Dewdney, 1992	To determine the impact of training in communications on reference interviews	Recorded and transcribed reference interviews	Quantitative	E. Z. Jennerich's Interviewing Skills Rating Scale (1974); modified to 4-point scale for each skill
Dilevko & Gottlieb, 2004	To determine the portrayal of librarians in popular culture	Obituaries in the *New York Times*	Qualitative	
Green, 1991	To identify conceptual models of information	Abstracts in the *LISA* database	Qualitative	
Haas & Grams, 1998a	To analyze Web pages and the links they contain to develop a classification system for both	1,500 links contained within 75 Web pages	Qualitative	
Haas & Grams, 1998b	To analyze links in Web pages and develop a taxonomy of functions 75 Web pages and 1,500 links contained therein	75 Web pages and 1,500 links contained therein.	Qualitative	
Haas & Grams, 2000	To discuss four overarching questions arising from past two studies of Web pages and links	75 Web pages and 1,500 links contained therein	Qualitative	
Hahn, 1999	To investigate how authors, editors and readers viewed development of electronic journals; to identify how authors decided to become involved in electronic publishing and how social structures influenced the process	Recorded and transcribed interviews	Qualitative	

Table 1. Selected Examples of Content Analysis in LIS Research, 1991–2005

Article[a]	Purpose	Data	Type of Analysis[b]	External Coding Scheme
Kracker & Wang, 2002	To investigate students' perceptions of research and research paper anxiety	Student writings on research paper experience	Qualitative/ Quantitive	
Lynch & Smith, 2001	To identify how computer technology is changing academic library positions	Job advertisements in *College & Research Libraries News*	Quantitative/ Qualitative	Yuan's (1996) Checklist of Computer-related Codes
Maloney-Krichmar & Preece, 2005	To develop an in-depth understanding of the dynamics of online group interaction and relationship between individual participants' online and offline lives	Messages on an electronic healthsupport list; interviews with participants.[c]	Qualitative/ Quantitative	Bales Transactional Analysis Schema (1951) Benne and Sheats's (1948) Group Membership Role Classification (1948)
Marsh & White, 2003	To develop a thesaurus of image-text relationships	Published research articles and books	Qualitative/ Quantitative	
Nitecki, 1993	To identify conceptual models of libraries among three groups	Letters and opinions in *Chronicle of Higher Education*	Qualitative	
Schoch & White, 1997	To compare communication patterns of participants in consumer health electronic lists for a chronic and an acute disease	Messages on two consumer health electronic lists	Quantitative/ Qualitative	Bales Transactional Analysis Schema (1951)
Stansbury, 2002	To determine the nature of problem statements in LIS articles	Research-based articles in eight core LIS journals	Qualitative/ Quantitative	Hernon & Metoyer-Duran (1993) attributes of problem statements
Wang & Gao, 2004	To analyze the extent and nature of technical services–oriented Web pages in academic libraries; to determine variations by type of institution	Technical services Web pages at U.S. academic libraries	Qualitative	
Wang & Soergel, 1998	To explore how real users select documents for their research projects from bibliographic searches	Interviews with participants as they made judgments about items in an online search	Qualitative	
Wang & White, 1999	To determine how real users make relevance judgments during use and citation phases of research projects	Interviews with participants about decision factors for using and citing documents in a research project	Qualitative	Wang & Soergel's (1998) Criteria for Document Selection (preliminary; modified substantially during coding)

Table 1. Selected Examples of Content Analysis in LIS Research, 1991–2005

Article[a]	Purpose	Data	Type of Analysis[b]	External Coding Scheme
White, 1998	To characterize questioning behavior in reference interviews preceding delegated online searches and to relate it to questioning behavior in other interviews/settings	Reference interviews	Qualitative/ Quantitative	Graesser's Typology of Questions (Graesser, McMahen & Johnson, 1994); White's (1985) Typology of Reference Interview Content (modified to add categories)
White, 2000	To characterize questioning behavior on a consumer health electronic list	Messages on an electronic list	Qualitative/ Quantitative	Graesser's Typology of Questions (Graesser, McMahen & Johnson, 1994); Roter's (1984) Typology of Medical Content of Patient Questions
White, Abels & Gordon-Murnane, 1998	To characterize adopters and non-adopters of Web for e-commerce	Web sites of publishers of business information	Qualitative	
White, Abels, & Agresta (in process)	To assess the relationship between chat interview quality and response quality	Transcripts of chat interviews	Quantitative	White, Abels & Kaske Typology of Turn Functions in Reference Interviews (2003)
White & Iivonen, 2001	To identify the reasons for selecting initial strategy in Web searches	Brief responses to questionnaire question about reasoning for decision	Qualitative	

[a]Complete references for items referred to in the first and last columns are in the bibliography.
[b]Studies are sometimes hybrids, with characteristics predominant to one type of content analysis but with some from the other type. For these, both types are sometimes noted with the predominant form first.
[c]Authors used other data in broader project; only data covered in this article or analyzed by content analysis are mentioned here.

matter) to the contexts of their use" (Krippendorff, 2004, p. 18). The notion of inference is especially important in content analysis. The researcher uses *analytical constructs,* or rules of inference, to move from the text to the answers to the research questions. The two domains, the texts and the context, are logically independent, and the researcher draws conclusions from one independent domain (the texts) to the other (the context). In LIS studies the analytical constructs are not always explicit.

The analytical constructs may be derived from (1) existing theories or practices; (2) the experience or knowledge of experts; and (3) previous research (Krippendorff, 2004, p. 173). Mayring (2000), the author of a standard German-language text on qualitative content analysis, suggests using a model of communication to determine the focal point for the inferences. Conclusions can be drawn about the communicator, the message or text, the situation surrounding its creation—including the sociocultural background of the communication—and/or the effect of the message. For example, Nitecki (1993) focuses on characterizing the communicator. She draws inferences about academicians' conceptual models of libraries based on analyzing the metaphors they used when they referred to libraries in published letters to the editor and opinion articles.

Content analysis involves specialized procedures that, at least in quantitative content analysis, allow for replication. The findings of a good study using quantitative content analysis, therefore, do not rely solely on the authority of the researchers doing the content analysis for their acceptability. They can be subjected to independent tests and techniques for judging their validity and reliability. Indeed, the extent to which validity and reliability can be judged are significant issues in evaluating a research methodology, and they are considered in subsequent sections in relation to both quantitative and qualitative content analysis.

Data

What constitutes data that can be used for content analysis studies? Most important is that the data provide useful evidence for testing hypotheses or answering research questions. Another key factor is that the data communicate; they convey a message from a sender to a receiver. Krippendorff's definition expands text to include "other meaningful matter" (2004, p. 18). Pictures on a Web site, for example, are used to convey one or more meanings, often in combination with text (Marsh & White, 2003) and, as such, can be subjected to content analysis either by themselves or by looking at the relationships between images and text, as Marsh and White have done. Both Bell (2001) and Collier (2001) discuss the content analysis of visual images.

Beaugrande and Dressler (1981) suggest seven criteria for defining a text, which is the more common form of data for content analysis: cohesion, coherence, intentionality, acceptability, informativity, situationality,

and intertextuality. In other words, text appropriate for content analysis is composed of linguistic elements arranged in a linear sequence that follows rules of grammar and dependencies and uses devices such as recurrence, anaphora and cataphora, ellipsis, and conjunctions to cause the elements to "hang together" to create a message (cohesion). The text has meaning, often established through relationships or implicature that may not be linguistically evident, and draws on frameworks within the recipient for understanding (coherence). The writer or speaker of the text intends for it to convey meaning related to his attitude and purpose (intentionality). Conversely, recipients of the message understand the text as a message; they expect it to be useful or relevant (acceptability). The text may contain new or expected information, allowing for judgments about its quality of informing (informativity). The situation surrounding the text affects its production and determines what is appropriate for the situation and the culture (situationality). The text is often related to what precedes and follows it, as in a conversation (one interpretation of intertextuality), or is related to other similar texts, for example, others within a genre, such as transcripts of chat sessions (another meaning of intertextuality).

The texts used in the LIS studies in Table 1 vary significantly. Some are generated in connection with the immediate research project; other texts occur naturally in the conduct of normal activities and independent of the research project. The former include responses to open questions on questionnaires (Kracker & Wang, 2002; White & Iivonen, 2001, 2002) and interviews with participants (Buchwald, 2000; Hahn, 1999). The latter include reference interviews (Dewdney, 1992), segments of published articles and books (Green, 1991; Marsh & White, 2003; Nitecki, 1993), obituaries (Dilevko & Gottlieb, 2004), problem statements in published articles (Stansbury, 2002), job advertisements (Croneis & Henderson, 2002; Lynch & Smith, 2001), messages on electronic lists (Maloney-Krichmar & Preece, 2005; White, 2000), and Web pages (Haas & Grams, 1998a, 1998b, 2000; Wang & Gao, 2004). Some studies use a combination of the two. For example, Buchwald (2000) analyzed recorded and transcribed informant interviews, observation notes generated during the research, and existing group documents in studying Canada's Coalition for Public Information's role in the federal information policy-making process.

Neuendorf (2002) proposes a useful typology of texts that takes into consideration the number of participants and/or setting for the message: individual messaging, interpersonal and group messaging, organizational messaging, and mass messaging. Individual responses to an open question on a questionnaire or in an interview are examples of individual messaging; the objective of content analysis is usually to identify that person's perspective on the topic. Reference interviews are a form of dyadic, interpersonal communication (Dewdney, 1992). Messages on electronic lists (Schoch & White, 1997) offer an example of group messaging; the person sends the

message to the group, any member of which can reply. The objective, in this case, is to characterize the communications of the group. Technical services Web sites (Wang & Gao, 2004), often existing only on Intranets, are examples of organizational communication. Job advertisements in LIS journals (Croneis & Henderson, 2002) are examples of mass messaging.

All of these types of text can occur within various applied contexts. For example, within the context of consumer health communication, studying messages on consumer-oriented health electronic lists (informal, group messaging) can provide insights into information needs that are not satisfied through doctor-patient interviews (more formal, interpersonal, dyadic communication) (White, 2000). Analyzing job advertisements (Croneis & Henderson, 2002) is similar to studying personal ads in other fields (Cicerello & Sheehan, 1995).

Data: Unitizing

At an early point in a content analysis study, the data need to be "chunked," that is, broken into units for sampling, collecting, and analysis and reporting. *Sampling units* serve to identify the population and establish the basis for sampling. *Data collection units* are the units for measuring variables. *Units of analysis* are the basis for reporting analyses. These units may be, but are not necessarily, the same. In many cases, the sampling unit is the documentary container for the data collection unit and/or units of analysis. It is the naturally occurring vehicle that can be identified and retrieved. In Dewdney (1992), for example, the entire interview serves as all three units. In White (2000) the message is the sampling unit; she has several different units of analysis in her study of questions in electronic lists: the message as a whole and individual questions within the messages. She also breaks the questions down into the form and content of the question, focusing on different segments of the question as phrased for categorizing.

In separate studies, Green (1991) and Nitecki (1993) focus on two words (*information* and the stem *librar,* respectively) and analyze the phrase immediately surrounding each occurrence of the word (data collection units) in two types of documents (sampling units) (for Green, abstracts in the *LISA* database; for Nitecki, letters and opinion articles in the *Chronicle of Higher Education*) to identify the metaphors surrounding use of these terms. They subsequently analyze the phrases to generate atomized phrases and then collapse them into metaphors (the unit of analysis). Each then interprets the metaphors as evidence of conceptual models held by the writers of the documents. In comparison to Dewdney (1992), who also studied reference interviews, White, Abels, and Agresta (2004) analyze turns (the unit of analysis) within chat reference interviews (the sampling unit). In Marsh and White (2003) the emphasis is on relationships between images and text, so the unit of analysis is the image-text pair, defined as the image and its related text segment (p. 652).

Pragmatism determines the sampling and data collection unit; the research question or hypothesis determines the unit of analysis. In all of the studies mentioned above, the unit of analysis is naturally related to the research question or hypothesis being addressed.

Procedures: Quantitative Content Analysis

Before discussing distinctions between qualitative and quantitative content analysis, it is useful to identify and explain the steps involved in content analysis. The focus initially is on the steps for a study using quantitative content analysis. The steps are as follows:

1. Establish hypothesis or hypotheses
2. Identify appropriate data (text or other communicative material)
3. Determine sampling method and sampling unit
4. Draw sample
5. Establish data collection unit and unit of analysis
6. Establish coding scheme that allows for testing hypothesis
7. Code data
8. Check for reliability of coding and adjust coding process if necessary
9. Analyze coded data, applying appropriate statistical test(s)
10. Write up results

Generating Hypotheses

Quantitative content analysis flows from a positivist research tradition and is deductive in its approach. Its objective is to test hypotheses, not to develop them. Drawing on related research and existing, relevant theory, a researcher first establishes one or more hypotheses that can be tested using content analysis. These hypotheses flow from what is already known about the problem and the extant research questions. In Dewdney, for example, "the hypothesis predicted, essentially, that interviews conducted by librarians who had received training in either neutral questioning or in the use of microskills would contain more examples of effective use of the skills taught, respectively, than interviews conducted by these same librarians before training, or than interviews conducted by librarians who had received no direct training" (1992, p. 131).

Determining Data for Analysis

The hypotheses, in turn, serve to guide subsequent decisions in the methodology. For example, they determine the nature of the data that would be required to test the hypotheses. In Dewdney (1992) it is clear that, to test her hypothesis, she needs to collect reference interviews under different situations: from librarians with training (1) before and (2) after the training, and (3) from librarians with no direct training.

Sampling

A major objective of social science research is generalizability, that is, the ability to generalize from the specific to the general—for example, to study the sample but infer from the sample's findings something about the population from which the sample is drawn. With a relatively nonstratified population, the ideal is random sampling, that is, sampling in which the probability of any unit within the population being selected is the same. To do this effectively, it is essential to know all units that exist in the population, such as all research articles published during a particular time period within a set of journals (Stansbury, 2002). Sometimes it is not possible to know all units beforehand, but a list can be generated as the sample is drawn. For example, to obtain a representative sample, randomly selected, from messages on two electronic lists and to ensure that the sampling period was sufficiently long to allow for getting a range of topics, messages, and participants, Schoch and White (1997) first did a preliminary study, based on archives of the lists, to determine the rate of messaging per list, or the average number of messages per month. At the start of the data-gathering period, all messages were downloaded and numbered separately for each list, and a sample of 1,000 messages was randomly chosen from the first 3,000 messages on each list written from the onset of the data-gathering period. Based on the messaging rate, the data-gathering period should have lasted approximately two months, but, because the rate of messaging actually varied across the two lists, data-collecting continued slightly longer in one list than in the other to achieve the same number of messages per list.

Coding

In quantitative content analysis the coding scheme is determined a priori, that is, before coding begins. A coding scheme operationalizes concepts that may in themselves be amorphous. It establishes categories that are relevant and valid. Relevant means that they allow for testing the hypotheses. Validity refers to "the extent to which a measuring procedure represents the intended, and only the intended, concept" (Neuendorf, 2002, p. 112). Validity can be assessed in several ways. Face validity, which is common in content analysis, refers to the extent to which a measure "gets at" the essential aspects of the concept being measured. Face validity is inherently subjective. To determine face validity, researchers assess as objectively as possible the correspondence between what they measure and how they measure it. One way of corroborating face validity is to have judges work backwards from the measure to determine the concept being measured (Neuendorf, 2002, p. 115). Other means of assessment are criterion validity, which relies on assessing the correspondence between the code and criteria, such as concurrent or predictive behavior or norms of behavior; content validity, which looks at the completeness of representation of the concept; and construct validity, which refers to "the extent to which

a measure is related to other measures (constructs) in a way consistent with hypotheses derived from theory" (Neuendorf, 2002, p. 117). Construct validity is more difficult to assess than criterion or content validity but is a worthy objective.

In addition, a good coding scheme has categories or levels that are exhaustive, that is, all relevant aspects of the construct are represented, are mutually exclusive, and are measured at the highest possible scale of measurement based on the four scales of measurement (nominal, ordinal, interval, and ratio).[4] The coding scheme should have clear definitions, easy-to-follow instructions, and unambiguous examples. All of these features promote the reliability of the coding, that is, the likelihood that all coders will code the same item the same way or that a coder will code the same item the same way at different points in time.[5] (For examples of coding schemes, see Haas & Grams, 2000, pp. 191–192; Hahn, 1999, Appendix B, pp. 229–237; Kracker & Wang, 2002, Appendices A-C, pp. 304–305; and Marsh & White, 2003, pp. 666–672.) If the coding scheme is modified during the coding, it must be re-applied to the data already coded so that all data are coded according to the same coding scheme.

The complexity of the coding scheme varies, and individual codes may be combined after the coding to develop a composite measurement, such as an index, or otherwise grouped to show relationships among the measures. Kracker and Wang (2002), for example, initially identified affective words that expressed emotions and subsequently clustered the categories into an affective classification scheme indicating negative and positive clusters for three major areas. Marsh and White (2003) grouped the image-text relationships into three categories: functions expressing little relation to the text; functions expressing close relation to the text; and functions going beyond the text.

Many content analysis studies do not develop their own coding scheme but rely instead on coding schemes devised by other researchers. Stansbury (2002) used the problem statement attributes identified by Hernon and Metoyer-Duran (1993) as a code for analyzing problem statements in LIS journals. Maloney-Krichmar and Preece (2005) and Schoch and White (1997) used Bales's Transactional Analysis Schema (Bales, 1951) to analyze messages on consumer health electronic lists. Using the same coding scheme across studies allows for easy comparisons among the studies. For example, after applying Graesser's Typology of Questions (Graesser, McMahen, & Johnson, 1994) to questions in reference interviews, White (1998) compared the incidence of questions and types of questions in reference interviews with similar question incidence data in tutoring sessions and decision support systems. In another study (White, 2000) coding the content of questions on consumer-health electronic lists with Roter's (1984) typology of questions in physician-patient interactions allowed for comparisons across the two settings. The last column in Table 1 shows

the content analytic schemes from other researchers used in quantitative content analysis studies.

Several coding schemes developed by LIS researchers have potentially broad use in LIS: (1) Haas and Grams' (1998a, 1998b, 2000) taxonomies for Web pages and links; (2) the two sets of categories developed by Kracker and Wang (2002) reflecting affective and cognitive aspects of Kuhlthau's (1993) Information Search Process (ISP) model; and (3) Marsh and White's (2003) taxonomy for analyzing text-image relationships in a variety of settings.

Just because coding schemes are developed a priori does not mean that the instances of the categories become immediately obvious and, as a result, easy to code. As in qualitative content analysis, the analysis often requires careful, iterative reading of the text. Marsh and White (2003) include several examples of image-text pairs, their codes, and the thinking surrounding coding each pair with their taxonomy of image-text relationships. These examples illustrate the complexity and depth of thinking that may be necessary in coding, even with an a priori coding scheme.

Analyzing the Coded Data

After the coding, which in itself is analytical, the researcher undertakes several additional steps. These steps, too, are done within the framework of the hypotheses or research questions. First, he[6] summarizes the findings identified during the coding, formulating and restating them so that they can be understood easily and are applicable to his hypotheses or research questions. Second, he identifies and articulates the patterns and relationships among his findings so that he can test his hypotheses or answer his research questions. Finally, he relates these more involved findings to those in other situations or other studies. The last step allows him to put his findings into perspective.

In the analysis, the content analyst chooses from among a variety of statistical approaches or techniques for presenting and testing his findings. They range in complexity and demands for different scales of measurement for the variables. The approach he selects takes into consideration not only the questions he is addressing but also the nature of the data and may include tabulations; cross-tabulations, associations, and correlations; multivariate techniques, such as multiple regression analysis; factor analysis and multidimensional scaling; images, portrayals, semantic nodes, and profiles; contingencies and contingency analysis; and clustering. Often, decisions about using these techniques are made in the planning phase of the project since they influence and build on decisions that, of necessity, must occur earlier in the project, such as establishing the level of measurement for a particular variable. The output of these techniques can be presented, in most cases, both in tabular and graphic form. Not all of these techniques are used in the LIS content analysis studies in Table 1. Tabulations, cross-tabulations, associations, and correlations are common (see, for example,

Schoch & White, 1997; Stansbury, 2002; White, 1998). White, Abels, and Gordon-Murnane (1998) use clustering techniques to develop a typology of innovators in a study of the content of publishers' Web sites and use it to profile publishers along a spectrum from traditionalist to innovator.

Procedures: Qualitative Content Analysis

Proponents of qualitative and quantitative content analysis often emphasize their differences, yet many similarities exist as well. Noting four common elements, Krippendorff, who covers both variants in his text, points out "the proponents of both approaches: [1] sample text, in the sense of selecting what is relevant; [2] unitize text, in the sense of distinguishing words or propositions and using quotes or examples; [3] contextualize what they are reading in light of what they know about the circumstances surrounding the text; and [4] have specific research questions in mind" (2004, p. 87). Table 2 characterizes the two types of content analysis along several dimensions. The most significant differences are the foci of this section.

Formulating Research Questions

In contrast with quantitative content analysis, qualitative content analysis flows from a humanistic, not a positivistic, tradition. It is inductive. Qualitative content analysis may yield testable hypotheses but that is not its immediate purpose. Replacing the hypotheses are foreshadowing questions, that is, open questions that guide the research and influence the data that are gathered. In qualitative content analysis, however, the text plays a slightly different role in that, as the researcher reads through the data and scrutinizes them closely to identify concepts and patterns, some patterns and concepts may emerge that were not foreshadowed but that are, nevertheless, important aspects to consider. In that case, the researcher may legitimately alter his interests and research questions to pursue these new patterns. For example, in Hahn's study of the author and editor as early adopters of electronic journals, she initially had three open, foreshadowing research questions, based, to some extent, on diffusion theory (Rogers, 1995): "1) How do authors and editors working closely with an electronic journal perceive electronic journals?; 2) What is the decision process that authors are using to decide to publish in an electronic journal?; 3) How do social factors influence the adoption decision?" (Hahn, 1999, p. 6). As her coding and analysis evolved, she added: "4) What key relations between the scientific community and the publishing system are affected by electronic publishing?" (p. 122). Krippendorff refers to this iterative process of "recontextualizing, reinterpreting, and redefining the research until some kind of satisfactory interpretation is reached" (2004, pp. 87–88) as a *hermeneutic loop*. This procedure may actually occur in quantitative content analysis studies but only at the development phase of the research design; the development phase is followed by adherence to the practices specified earlier.

Table 2. Characteristics of Quantitative and Qualitative Content Analysis

Category	Quantitative	Qualitative
Research approach	Deductive; based on previous research, which allows for formulating hypotheses about relationships among variables	Inductive; research questions guide data gathering and analysis but potential themes and other questions may arise from careful reading of data
Research tradition or orientation	Positivist	Naturalist or humanist; hermeneutics
Objective	To make "replicable and valid inferences from texts . . . to the contexts of their use" (Krippendorff, 2004, p. 19)	"To capture the meanings, emphasis, and themes of messages and to understand the organization and process of how they are presented" (Altheide, 1996, p. 33); "Search for multiple interpretations by considering diverse voices (readers), alternative perspectives (from different ideological positions), oppositional readings (critiques), or varied uses of the texts examined (by different groups)" (Krippendorff, 2004, p. 88)
Data: Nature	Syntactic, semantic, or pragmatic categories; naturally occurring texts or text generated for project	Syntactic, semantic, or pragmatic categories; naturally occurring texts or text generated for project
Data: Selection	Systematic, preferably random, sampling to allow for generalization to broader population; data selection usually complete prior to coding	Purposive sampling to allow for identifying complete, accurate answers to research questions and presenting the big picture; selection of data may continue throughout the project
Categorization schema	Coding scheme developed a priori in accord with testing hypotheses; if adjustments are made during coding, items already coded must be recoded with the revised scheme; may use coding scheme(s) from other studies	Coding scheme usually developed in the process of close, iterative reading to identify significant concepts and patterns
Coding	Objective; tests for reliability and validity	Subjective; in some cases, use of memos to document perceptions and formulations; techniques for increasing credibility, transferability, dependability, and confirmability of findings

Table 2. Characteristics of Quantitative and Qualitative Content Analysis

Category	Quantitative	Qualitative
Argument basis for proof	Frequency, indicating existence, intensity, and relative importance; data allow for statistical testing of hypotheses; objectives are usually to generalize to broader population and to predict; interpretations may be supported by quotations from text	Deep grounding in the data; if numbers are presented, they are usually presented as counts and percentages; description of specific situation or case accurately and thoroughly; may involve triangulation based on multiple data sources for same concept; may use techniques to develop grounded theory to relate concepts and to suggest hypotheses that can be tested deductively; presentation "Support[s] interpretations by weaving quotes from the analyzed texts and literature about the contexts of those texts into their conclusions, by constructing parallelisms, by engaging in triangulations, and by elaborating on any metaphors they can identify" (Krippendorff, 2004, p. 88)
Use of computers	For dictionary-based content analysis or for developing environments prior to dictionary-based content analysis; also for statistical tests; representative software for content analysis: VBPro, WordStat	As annotation and searching aids; representative software: Atlas. TI or NVivo

Sampling

Both qualitative and quantitative content analysis researchers sample text and choose text that is relevant for their purpose, but qualitative researchers focus on the uniqueness of the text and are consciously aware of the multiple interpretations than can arise from a close perusal of it. The need for close, reiterative analysis itself usually limits the size of the sample.

In addition, since the object of qualitative research is not generalizability but transferability, sampling does not need to insure that all objects being analyzed have an equal or predictable probability of being included in the sample. Transferability refers to a judgment about whether findings from one context are applicable to another. Instead, the sampling should be theoretical and purposive. It may have as its objective providing the basis for identifying all relevant patterns in the data or characterizing a phenomenon. It may even present the findings quantitatively through numbers

and percentages but not through inferential statistics. Some cases may be selected prior to initiating coding, but the selection and coding may also occur in tandem, with subsequent case selection influenced by discoveries during the coding process. Analyzing new cases may continue until no new patterns or findings related to the concept under analysis become apparent in the coding process. If no new patterns are being found, generally the presumption is that all relevant patterns have been discovered and additional work would only confirm that finding. If at this point there is interest in noting the prevalence of a particular pattern, the researcher may move to using the pattern or range of patterns as a coding scheme and analyzing a body of documents. But, because the sampling is purposive, the researcher cannot extrapolate from the sample to the population.

Coding

For qualitative coding, the researcher's initial foci are not a priori codes but the initial foreshadowing questions he aims to answer through his research. The questions guide his initial approach to the data, but the process is inductive, not deductive. The evidence plays almost as significant a role in shaping the analysis as the initial questions. It is not unusual to have a person doing qualitative content analysis read through the data initially with the intent of trying to see the big picture. As he reads through the documents, he begins to tag key phrases and text segments that correspond to those questions, notes others that seem important but are unexpected, sees similarities in expressing the same concept, and continues iteratively to compare the categories and constructs that emerge through this process with other data and re-reading of the same documents. In the process, he may be looking for diversity of ideas, alternative perspectives, oppositional writings, and/or different uses of the texts, perhaps by different groups.

Data collection units and units of analysis vary. The researcher continually checks his growing interpretation of answers to his research questions against the documents and notes, especially situations that do not fit the interpretation or suggest new connections. In this way, he looks not only at confirming evidence of his emerging construct(s) but also at disconfirming evidence that needs to be considered as he presents his case for his interpretation. The overall process may suggest new questions that were not anticipated at the start of the analysis. Glaser and Strauss (1967) refer to the *constant comparison* approach to data analysis, in which the emerging relationships and categories are continually refined and emerging theory or patterns tested as new data are compared with old (see also Boeije, 2002).

To keep track of the developing concepts and the models that are emerging about how the concepts relate to each other, the researcher records his decisions and comments in memos. Two types of memos are common: concept memos, which logically focus on emerging concepts, the distinctive ways in which these are phrased, and his own interpretation of the con-

cepts; and theory memos, in which he focuses on relationships among the concepts and gradually integrates these concepts into a workable model. Memos reveal the subtleties of the researcher's interpretation and understanding of the constructs over time. In a conceptual memo, for example, Hahn (1998) comments:

> Thinking over some of the features of discussions that I feel are recurring but not previously captured by existing coding structures, I initially considered the concept of advantages and disadvantages. However, it seems like a more useful organizing conceptual structure is one of optimizing characteristics. The idea is that these are characteristics of the journal perceived by the community. The editors and publishers try to optimize these to encourage both submissions and readership. Authors also try to make an optimal match with these characteristics given the nature of the paper they have in hand ready for submission. (n.p.)

Qualitative content analysis has developed approaches similar to validity and reliability for assessing the rigor of the coding and analysis process. Qualitative content analysis focuses on creating a picture of a given phenomenon that is always embedded within a particular context, not on describing reality objectively. Lincoln and Guba (1985) describe four criteria used to assess the degree to which a qualitative study will have "truth value," that is, "confidence in the 'truth' of the findings of a particular inquiry" (Guba & Lincoln, 1981, p. 246): credibility, transferability, dependability, and confirmability.[7] Credibility, the equivalent of internal validity, calls for identifying all important factors in the research question and accurately and completely describing the ways in which these factors are reflected in the data gathered. Transferability, or external validity, is essentially a judgment about the applicability of findings from one context to another. Generally a qualitative researcher tries to situate his findings within a relevant theoretical paradigm, understanding that findings sensible within it can be applied to other, comparable contexts with greater confidence. Similarly, he usually tries to collect data on a single factor or question aspects from multiple sources with the understanding that findings based on multiple data sources can be transferred with greater confidence. Collecting, analyzing, and cross-checking a variety of data on a single factor or aspect of a question from multiple sources, and perhaps perspectives, as Buchwald (2000) did, is termed *triangulation* and is a way to heighten a qualitative study's credibility and confirmability.

Dependability addresses the notion of replicability and defines it as "stability after discounting . . . conscious and unpredictable (but rational and logical) changes" (Guba & Lincoln, 1981, p. 247) in findings during repetitions of the study. Confirmability relates to objectivity and is measured in quantitative content analysis by assessing inter-rater reliability. In qualitative research findings are confirmed by looking at the data, not the researcher(s), to determine if the data support the conclusions. The

important criterion is not numeric correspondence between coders but conceptual consistency between observation and conclusion.

Method of Analysis

Analysis is integrated into coding much more in qualitative content analysis than in quantitative content analysis. The emphasis is always on answering the research questions but considering as well any transformations that the initial foreshadowing questions may have undergone during the coding or any new questions or themes that emerge during the coding. Often the result of qualitative analysis is a composite picture of the phenomenon being studied. The picture carefully incorporates the context, including the population, the situation(s), and the theoretical construct. The goal is to depict the "big picture" of a given subject, displaying conceptual depth through thoughtful arrangement of a wealth of detailed observations.

In presenting the results the researcher may use numbers and/or percentages, either in simple tabulations or in cross-tabulations to show relationships, but he may also rely simply on the gradual accretion of details within his textual presentation without resort to numbers. Often the analysis results in both graphic and tabular presentation of models elicited during the analysis. Wang and White (1999), for example, present a graphic model of document use at three different stages in a research project, showing the criteria and decision rules the researchers applied at each stage (see Figure 6, p. 109). This table incorporates data from a previous study, which covered the first stage (Wang & Soergel, 1998), and is supported in the second study by data in Tables 2 and 4 (Wang & White 1999, pp. 104, 107, respectively) for criteria and decision rules in the second and third stages, respectively. The tables present, for each criterion and decision rule, the number of users mentioning each and the number of documents about which they were mentioned.

The text may be a narrative of findings about the phenomenon being studied with quotations to illustrate the conclusions. In the same study, for example, the authors refer to the participants' use of reputation as a criterion in determining relevance:

> Participants comment on whether or not the document is written by a reputable author or organization or published in a reputable journal. An example related to the document's author is "It is by a very minor person, X; Y [co-author] is a little better known than X. I know them by reputation. I don't know them personally." Another example comments on the authority of the publisher or the author's affiliation: "I was looking for something which wouldn't have a bias. The World Bank is accepted by all countries. We already know that the World Bank is very involved in sending technical support or funding for such projects" (Wang & White, 1999, p. 105).

Ahuvia (2001) suggests that reviewers can better judge the confirmability or public credibility of a qualitative content analysis if the researcher submits his original data set, codings, and justification for particular codes if necessary along with a manuscript. In a published study, the data, or at least a random subset, can be included as an appendix.

Using Computer Software

Depending on the number of documents, content analysis can be tedious and benefits enormously from the use of computers for a variety of tasks. Collectively, the software programs serve in several capacities:

- As a research assistant, making it easy to markup the data, divide them into chunks for analysis, write notes, group together multiple instances of the same classification, and allow for global editing and coding.
- As a manipulator and extractor of data, matching the text against specialized dictionaries for coding purposes.
- As data collections, maintaining the electronic and coded versions, keeping track of all steps in the analysis, and, in the latter case, allowing for replicating the analysis.
- As a means for doing or facilitating quantitative analyses, such as frequency counts and percentages, either within the program itself or by exporting data to statistical packages, thereby eliminating errors that would occur in multiple inputs of the data. The statistical packages would usually allow for inferential statistics. (Mayring, 2000)

The programs arrange themselves on a spectrum from simply facilitating a human's coding of the electronic data to direct involvement in analyzing the document; matching terms to an electronic dictionary, which is a coding scheme; and coding the data. In the latter human input occurs primarily in developing the dictionary and in interpreting the results of the coding. In the middle is a set of programs that facilitates developing the dictionaries used in the latter. Lowe (2002) refers to these respectively as annotation aids, developing environments, and programs for dictionary-based content analysis. Examples of the first are NVivo (2003–2005), QSR N6 (2005) and Atlas-TI (Muhr, 2004). These programs now allow for storing not only textual documents but also images and audio in electronic form. Qualitative content analysis relies more on annotation aids. Dictionary-based content analysis programs rely on several basic functions: word and category counts and frequency analyses, visualization (including clustering), and sometimes concordance generation. DIMAP-4 (Litkowski & McTavish, 2001) and KEDS (Schrodt, 1996), and TABARI (Schrodt, 2000) are examples of developing environments. WordStat 5.0 (Peladeau, 2005), VBPro (Miller, 2003), and the General Inquirer (Stone, 2002; The General Inquirer, n.d.) are examples of dictionary-based content analysis programs. LIS researchers do not always identify the software used in analyses. Agosto

and Hughes-Hassell (2005) mention NVivo; Marsh (2002) uses Atlas-TI; White (1998) and Kracker and Wang (2002) use QSR NUD*IST, renamed, in its latest version, as QSR N6.

Conclusion

Content analysis is a highly flexible research method that has been widely used in LIS studies with varying research goals and objectives. The research method is applied in qualitative, quantitative, and sometimes mixed modes of research frameworks and employs a wide range of analytical techniques to generate findings and put them into context. The LIS studies referred to in this article are not always purist but occasionally use a hybrid approach, incorporating elements of qualitative and quantitative content analysis for good reason. This article characterizes content analysis as a systematic, rigorous approach to analyzing documents obtained or generated in the course of research. It briefly describes the steps involved in content analysis, differentiates between quantitative and qualitative content analysis, and shows that content analysis serves the purposes of both quantitative research and qualitative research. In addition, the article serves as a gateway to selected LIS studies that have used content analysis and to methodological books and articles that provide more detail about aspects of content analysis discussed only briefly here.

Bibliographic Note

Two recent texts on content analysis are Krippendorff (2004) and Neuendorf (2002). Krippendorff covers both quantitative and qualitative content analysis; Neuendorf focuses on quantitative content analysis. Neuendorf (2005) maintains a text-related Web site with many useful resources: the Content Analysis Guidebook Online (http://academic.csuohio.edu/kneuendorf/content). Titscher, Meyer, Wodak, and Vetter (2000) provide chapters for specific types of textual analysis not covered in this article; Schiffrin (1994) discusses various types of discourse analysis. Additional useful methodological chapters are Bell (2001) and Collier (2001) for content analysis of visual images and Evans (2002) for dictionary-based content analysis.

Articles reviewing software are useful but become dated quickly; Skalski's (2002) review in Neuendorf's (2002) text has a tabular presentation of software features in addition to paragraphs describing about twenty programs; his table establishes a useful framework for evaluating software. Several Web sites maintain more current reviews and/or links to content analysis software publisher pages. See, for example, the "Classification of Text Analysis Software" section of Klein's (2002–2003) Text Analysis Info Page (http://www.textanalysis.info) and the content analysis resources listed under the software section of Evans's (n.d.) Content Analysis Resources (http://www.car.ua.edu). Krippendorff's (2004) chapter 12 on computer aids is also useful for showing how computers can aid content analysis.

The Web sites mentioned above (Neuendorf, Klein, and Evans) are the most useful for content analysis researchers. Contents analysis researchers in all fields communicate informally via the Content Analysis News and Discussion List (2006) (content@bama.ua.edu). Its archives are available at http://bama.ua.edu/archives/content.html.

Notes

The authors are grateful to Karla Hahn for permitting a quotation from a concept memo, Susan Davis for comments, and the authors whose works are mentioned in this article for their careful and clear presentation of their methodology.

1. Berelson's (1952) *Content Analysis in Communications Research* is considered the "first systematic presentation" of the conceptual and methodological elements of content analysis and "codified the field for years to come" (Krippendorff, 2004, p. 8).
2. For a useful discussion and explanation of each type, see Krippendorff (2004), Schiffrin (1994), and Titscher, Meyer, Wodak, and Vetter (2000). Titscher et al. includes a map of theories and methods that is notable for illustrating relationships among them (Figure 4.1, p. 51).
3. Krippendorff's (2004) text considers both quantitative and qualitative content analysis. Another recent text by Neuendorf (2002) focuses on quantitative content analysis.
4. Any statistics text should discuss scales of measurement. See, for example, StatSoft, Inc.'s (2004) *Electronic Statistics Textbook.*
5. See Lombard, Snyder-Duch, and Bracken's (2005) *Practical Resources for Assessing and Reporting Intercoder Reliability in Content Analysis Research Projects.* This paper is invaluable in discussing the reasons for assessing and reporting intercoder reliability, the proper steps involved in doing so, the preferred statistical tests, and the information to be reported, among other topics. Krippendorff (2004) also includes useful sections on reliability (chap. 11, pp. 211–256) and validity (chap. 13, pp. 313–338).
6. Throughout this article, when *he, his,* and *him* are used without the context of a specific researcher, they refer to researchers of both genders.
7. Lincoln and Guba (1985) apply these to qualitative research studies generally, not just to coding, but they are also applicable in the narrower context.

References

Agosto, D. E. & Hughes-Hassell, S. (2005). People, places, and questions: An investigation of the everyday life information-seeking behaviors of urban young adults. *Library & Information Science Research, 27,* 141–163.

Ahuvia, A. (2001). Traditional, interpretive, and reception based content analyses: Improving the ability of content analysis to address issues of pragmatic and theoretical concern. *Social Indicators Research, 54,* 139–172.

Altheide, D. L. (1996). *Qualitative media analysis.* Thousand Oaks, CA: Sage.

Bales, R. F. (1951). *Interaction process analysis: A method for the study of small groups.* Cambridge, MA: Addison-Wesley Press.

Beaugrande, R. D., & Dressler, W. U. (1981). *Einführung in die textlinguistik.* Tübingen: Niemeyer.

Bell, P. (2001). Content analysis of visual images. In T. Van Leeuwen & C. Jewitt (Eds.), *Handbook of visual analysis* (pp. 10–34). Thousand Oaks, CA: Sage.

Benne, K. D., & Sheats, P. (1948). Functional roles of group members. *Journal of Social Issues, 41,* 41–49.

Berelson, B. (1952). *Content analysis in communications research.* New York: Free Press.

Boeije, H. (2002). A purposeful approach to the constant comparison method in the analysis of qualitative interviews. *Quality & Quantity, 36,* 391–409.

Buchwald, C. C. (2000). A case study of Canada's Coalition for Public Information in the information highway policy-making process. *Library & Information Science Research, 22,* 123–144.

Cicerello, A., & Sheehan, E. P. (1995). Personal advertisements: A content analysis. *Journal of Social Behavior and Personality, 10,* 751–756.

Collier, M. (2001). Approaches to analysis in visual anthropology. In T. Van Leeuwen & C. Jewitt (Eds.), *Handbook of visual analysis* (pp. 35–60). Thousand Oaks, CA: Sage.

Content Analysis News and Discussion List. (2006). Archives of CONTENT@BAMA.UA.EDU. Retrieved March 1, 2006, from http://bama.ua.edu/archives/content.html.

Croneis, K., & Henderson, P. (2002). Electronic and digital librarian positions: A content analysis of announcements from 1990 through 2000. *Journal of Academic Librarianship, 28,* 232–237.

Dewdney, P. (1992). Recording the reference interview: A field experiment. In J. D. Glazier & R. R. Powell (Eds.), *Qualitative research in information management* (pp. 122–150). Englewood, CO: Libraries Unlimited.

Dilevko, J., & Gottlieb, L. (2004). The portrayal of librarians in obituaries at the end of the twentieth century. *Library Quarterly, 74*(2), 152–180.

Evans, W. (2002). Computer environments for content analysis: Reconceptualizing the role of humans and computers. In O. V. Burton (Ed.), *Computing in the social sciences and humanities* (pp. 67–86). Urbana, IL: University of Illinois Press.

Evans, W. (n.d.). *Content analysis resources.* Available at the University of Alabama. Retrieved March 1, 2006, from http://www.car.ua.edu.

The General Inquirer [computer software]. (n.d.). Princeton, NJ: The Gallup Organization. Retrieved November 2, 2005, from the WebUse Project at the University of Maryland Web site, http://www.webuse.umd.edu:9090.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine Press.

Graesser, A. C., McMahen, C. L., & Johnson, B. K. (1994). Question asking and answering. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 517–523). San Diego, CA: Academic Press.

Green, R. (1991). The profession's models of information: A cognitive linguistic analysis. *Journal of Documentation, 47,* 130–148.

Guba, E. G., & Lincoln, Y. S. (1981). Epistemological and methodological bases of natural inquiry. *Educational Communication & Technology Journal, 30*(4), 233–252.

Haas, S. W. & Grams, E. S. (1998a). Page and link classifications: Connecting diverse resources. In I. H. Witten, R. M. Akscyn, & F. M. Shipman (Eds.), *Proceedings of Digital Libraries '98—Third ACM Conference on Digital Libraries, June 23–26, 1998, Pittsburgh, PA* (pp. 99–107). New York: Association for Computing Machinery.

Haas, S. W. & Grams, E. S. (1998b). A link taxonomy for Web pages. In C. M. Preston (Ed.), *Information access in the global information economy, Proceedings of the 61st Annual Meeting of the American Society for Information Science, Pittsburgh, PA, October 26–29, 1998* (pp. 485–495). Medford, NJ: Information Today.

Haas, S. W., & Grams, E. S. (2000). Readers, authors, and page structure: A discussion of four questions arising from a content analysis of Web pages. *Journal of the American Society for Information Science, 51*(2), 181–192.

Hahn, K. (1998). *Memo: Optimizing characteristics of journal titles.* Unpublished memo provided to author, October 2005.

Hahn, K. (1999). *Electronic journals as innovations: A study of author and editor early adopters.* Unpublished doctoral dissertation, University of Maryland at College Park, Maryland.

Hernon, P., & Metoyer-Duran, C. (1993). Problem statements: An exploratory study of their function, significance, and form. *Library & Information Science Research, 15*(1), 71–92.

Jennerich, E. Z. (1974). *Microcounseling in library education.* Unpublished doctoral dissertation, University of Pittsburgh, 1974.

Klein, H. (2002–2003). *Welcome to text analysis info.* Rudolstat: Social Science Consulting. Retrieved November 2, 2005, from http://www.textanalysis.info.

Kracker, J., & Wang, P. (2002). Research anxiety and students' perceptions of research: An experiment. Part II. Content analysis of their writings on two experiences. *Journal of the American Society for Information Science & Technology, 53*(4), 294–307.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage.

Kuhlthau, C. (1993). *Seeking meaning: A process approach to library and information services.* Norwood, NJ: Ablex.

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

Litkowski, K., & McTavish, D. (2001). DIMAP-4 (DIctionary MAintenance Programs) [computer software]. Damascus, MD: CL Research.

Lombard, M., Snyder-Duch, J., & Bracken, C.C. (2005). *Practical resources for assessing and reporting intercoder reliability in content analysis research projects.* Retrieved May 27, 2005, from Temple University, School of Communications and Theater, Doctoral Program in Mass Media & Communication, http://www.temple.edu/mmc/reliability.

Lowe, W. (2002). *Software for computer analysis—A review.* Retrieved October 14, 2005, from Harvard University, The Institute for Quantitative Social Science, http://people.iq.harvard.edu/~wlowe/Publications/rev.pdf.

Lynch, B. P., & Smith, K. R. (2001). The changing nature of work in academic libraries. *College & Research Libraries, 62*(5), 407–420.

Maloney-Krichmar, D. & Preece, J. (2005). A multilevel analysis of sociability, usability, and community dynamics in an online health community. *ACM Transactions on Computer-Human Interaction, 12*(2), 201–232.

Marsh, E. E. (2002). *Rhetorical relationships between images and text in Web pages.* Unpublished doctoral dissertation, University of Maryland at College Park, Maryland.

Marsh, E. E., & White, M. D. (2003). A taxonomy of relationships between images and text. *Journal of Documentation, 59*(6), 647–672.

Mayring, P. (2000). Qualitative content analysis. *Forum Qualitative Social Research/Forum Qualitative Sozialforschung, 1*(2). Retrieved September 24, 2005, from http://www.qualitative-research.net/fqs-texte/2-00/2-00mayring-e.htm.

Miller, M. M. (2003). VBPro (Verbatim Protocol) [computer software]. Freeware.

Muhr, T. (2004). Atlas.ti5 [computer software]. Berlin: ATLAS.ti Scientific Software Development GmbH.

Neuendorf, K. A. (2002). *The content analysis guidebook.* Thousand Oaks, CA: Sage.

Neuendorf, K. A. (2005). *The content analysis guidebook online: An accompaniment to the content analysis guidebook by Kimberly A. Neuendorf.* Cleveland, OH: Cleveland State University. Retrieved March 1, 2006, from http://academic.csuohio.edu/kneuendorf/content.

Nitecki, D. A. (1993). Conceptual models of libraries held by faculty, administrators, and librarians: An exploration of communications in the *Chronicle of Higher Education*. *Journal of Documentation, 49*(3), 255–277.

NVivo 2.0 [computer software]. (2003–2005). Doncaster, Victoria, Australia: QSR International.

Peladeau, N. (2005). WordStat 5.0 [computer software]. Montreal: Provalis Research.

QSR N6 [computer software]. (2005). Durham, UK: QSR Software.

Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York: Free Press.

Roter, D. L. (1984). Patient question asking in physician-patient interaction. *Health Psychology, 3*(5), 395–409.

Schiffrin, D. (1994). *Approaches to discourse.* Cambridge, MA: Blackwell.

Schoch, N. A., & White, M. D. (1997). A study of the communications patterns of participants in consumer health electronic discussion groups. In C. Schwartz, & M. Rorvig (Eds.), *Digital collections, implications for users, funders, developers, and maintainers, Proceedings of the 60th Annual Meeting of the American Society for Information Science, Washington, DC, November 1–6, 1997* (pp. 280–292). Medford, NJ: Information Today.

Schrodt, P. A. (1996). KEDS (Kansas Event Data System) [computer software]. Lawrence: Department of Political Science, University of Kansas.

Schrodt, P. A. (2000). TABARI 0.5.1 (Textual Analysis By Augmented Replacement Instructions) [computer software]. Lawrence: Department of Political Science, University of Kansas.

Skalski, P. (2002). Computer content analysis software. In K. A. Neuendorf, *The content analysis guidebook* (pp. 225–235). Thousand Oaks, CA: Sage.

Stansbury, M. C. (2002). Problem statements in seven LIS journals: An application of the Hernon/Metoyer-Duran attributes. *Library & Information Science Research, 24*(2), 157–168.

StatSoft, Inc. (2004). *Electronic Statistics Textbook.* Tulsa, OK: StatSoft. Retrieved October 31, 2005, from StatSoft, Inc., http://www.statsoft.com/textbook/stathome.html.

Stone, P. J. (2002). *Welcome to the General Inquirer Home Page.* Cambridge, MA: Harvard College. Retrieved March 1, 2006, from http://www.wjh.harvard.edu/~inquirer.

Titscher, S., Meyer, M., Wodak, R., & Vetter, E. (2000). *Methods of text and discourse analysis* (B. Jenner, Trans.). Thousand Oaks, CA: Sage.

Wang, J., & Gao, V. (2004). Technical services on the net: Where are we now? A comparative

study of sixty Web sites of academic libraries. *Journal of Academic Librarianship, 30*(3), 218–221.

Wang, P., & Soergel, D. (1998). A cognitive model of document use during a research project. Study I. Document selection. *Journal of the American Society for Information Science, 49*(2), 115–133.

Wang, P., & White, M. D. (1999). A cognitive model of document use during a research project. Study II. Decisions at the reading and citing stages. *Journal of the American Society for Information Science, 50*(2), 98–114.

White, M. D. (1985). Evaluation of the reference interview. *RQ, 25,* 76-84.

White, M. D. (1998). Questions in reference interviews. *Journal of Documentation, 54*(4), 443–465.

White, M. D. (2000). Questioning behavior on a consumer health electronic list. *Library Quarterly, 70*(3), 302–334.

White, M. D., Abels, E. G., & Agresta, J. (2004). The relationship between interaction characteristics and answer quality in chat reference service. In *Online proceedings of the Virtual Reference Desk Conference, Cincinnati, OH, November 8–9, 2004.* Retrieved November 2, 2005, from http://www.vrd.org/conferences/VRD2004/proceedings/presentation.cfm?PID=376.

White, M. D., Abels, E. G., & Gordon-Murnane, L. (1998). What constitutes adoption of the Web: A methodological problem in assessing adoption of the World Wide Web for electronic commerce. In C. M. Preston (Ed.), *Information access in the global information economy, Proceedings of the 61st Annual Meeting of the American Society for Information Science, Pittsburgh, PA, October 26–29, 1998* (pp. 217–226). Medford, NJ: Information Today.

White, M. D., Abels, E. G., & Kaske, N. (2003). Evaluation of chat reference service quality: Pilot study. *D-LIB Magazine, 9.* Retrieved November 2, 2005, from http://www.dlib.org/dlib/february03/white/02white.html.

White, M. D., & Iivonen, M. (2001). Questions as a factor in Web search strategy. *Information Processing & Management, 37,* 721–740.

White, M. D., & Iivonen, M. (2002). Assessing the level of difficulty of Web search questions. *Library Quarterly, 72*(2), 207–233.

Yuan, Z. (1996). Analysis of trends in demand for computer-related skills for academic librarians from 1974 to 1994. *College & Research Libraries, 57,* 259–272.

Marilyn Domas White is an associate professor in the College of Information Studies at the University of Maryland, College Park, where she teaches in the area of information access. Her Ph.D. is in library science from the University of Illinois. Her current areas of research are information behavior, especially questioning behavior; information access, especially to electronic images; and scholarly communication. Recent publications include "A taxonomy of relationships between images and text," *Journal of Documentation, 59*(2003), 647–672, with Emily E. Marsh; "Evaluation of chat reference service quality: Pilot study," *D-Lib Magazine, 9*(2003), with Eileen G. Abels and Neal Kaske; and "Assessing the level of difficulty of Web search questions," *Library Quarterly, 72*(2002), 207–233, with Mirja Iivonen. Until recently, she was co-PI for the Computational Linguistics for Metadata Building (CLiMB-2) Project at the University of Maryland.

Emily E. Marsh is a consultant in information science and has been an adjunct faculty member at the College of Information Studies at the University of Maryland, College Park, where she teaches in the area of user behavior. Her Ph.D. is in library and information science from the University of Maryland (2002). Her areas of research interest are information design, illustration, and research methods. Recent publications include "A taxonomy of relationships between images and text," *Journal of Documentation, 59*(2003), 647–672, with Marilyn Domas White.

The Critical Incident Technique and the Qualitative Evaluation of the Connecting Libraries and Schools Project

Marie L. Radford

Abstract

This article describes Flanagan's Critical Incident Technique (CIT) for those seeking guidance in its application. Examples in the library and information science (LIS) field are discussed, including an in-depth example of a CIT study conducted as part of a qualitative evaluation of the Connecting Libraries and Schools Project (CLASP) in New York City. The CLASP study analyzed critical incidents from 2,416 fifth and seventh grade students regarding their perceptions of interactions with urban public librarians and library staff. For both positive and negative critical incidents, the most important factor in these preadolescent's perception of successful library visits is the attitude of the librarian or staff member they encounter.

Introduction

This article discusses Flanagan's (1954) Critical Incident Technique (CIT) and describes its applications in the library and information science (LIS) field. The CIT is a qualitative method designed to draw out the most memorable aspects of an event or experience from the study's participants (Ruben 1993). It has been used to evaluate programs or services and to inform their improvement. CIT questions typically have this format: Remember a time when you had a successful (specify activity)? Please describe. What was it about (specified activity) that made it successful? Or the negative: Remember a time when you had an unsuccessful (specify activity)?

Guidance is provided by this article for those considering using CIT in their research by discussing an example of a CIT study conducted by the author as part of a qualitative evaluation of the Connecting Libraries and

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 46–64
© 2006 The Board of Trustees, University of Illinois

Schools Project (CLASP). CLASP, a citywide program of The New York Public Library, the Brooklyn Public Library, and the Queens Borough Public Library, was evaluated by collecting critical incidents from questionnaires completed by 2,416 fifth and seventh grade students to gather their perceptions of encounters with public librarians.

Using the Critical Incident Technique in a Library Setting

CLASP was created by The New York Public Library in 1991 through a grant from the DeWitt Wallace-Reader's Digest Fund and was expanded to all five Manhattan boroughs in 1994 (Tice, 2001). The author was asked to conduct a qualitative evaluation of the impact of CLASP on students and chose the CIT as the most suitable method. Gathering students' critical incidents allowed analysis to center on how effective CLASP had been in changing or improving student's attitudes toward the public library.[1] The study had a twofold purpose: (1) evaluating CLASP, and (2) advancing knowledge of preadolescent perceptions of librarian/library staff encounters in the urban public library setting.[2] The CIT was selected as an appropriate method for achieving these purposes.

Flanagan's (1954) seminal article on the CIT documents the technique's origins and provides guidelines for its use. According to Flanagan, the CIT has five key stages, each of which will be discussed below with examples of its application in the CLASP evaluation:

1. General aims
2. Plans and specifications
3. Collecting the data
4. Analyzing the data
5. Interpreting and reporting

Stage 1: General Aims

During this stage, the key decision is to determine the purpose of the study: What does the researcher want to find out (Redman, Lambrecht, & Stitt-Gohdes, 2000)? Also, the researcher needs to conduct a literature review to discover what is already known about the type of activity, program, or participants to be studied. The CIT is useful for evaluating particular activities such as conducting outcomes assessments (Bycio & Allen, 2004; Jacoby & O'Brien, 2005) and measuring customer satisfaction (Arnold, Reynolds, & Ponder, 2005). Specifically, the researcher using CIT is looking to find out "precisely what it is necessary to do and not to do if the activity is to be judged successful or effective" (Fisher & Oulton, 1999, p. 113). In the case of the CLASP project, the study's purpose was program evaluation.

Stage 2: Plans and Specifications

Before any data can be collected, sufficient preparation is necessary. One important decision is identifying the subjects from whom critical incidents will be collected. In the case of the CLASP study, fifth and seventh graders, aged nine to twelve years old, were targeted because they would have experienced several CLASP classroom visits and they would have the writing skills necessary to complete the questionnaires. This highlights another decision to be made—the method of data collection. Critical incidents can be collected by observation, face-to-face (individual or group) interviews, telephone interviews, or paper or email questionnaires. This decision will be influenced by budgetary and staff considerations. In the CLASP study, project librarians were available to administer the paper questionnaires at the various school sites, to distribute and collect the surveys, and to assist students as needed.

Data collection through observation or face-to-face, group, or telephone interviews is much more labor-intensive when compared to paper or email survey distribution, but it has the advantage of providing the ability to ask probing or clarification questions of the subjects. This leads into the next decision point—determining who will collect the data. For a local project, such as the CLASP review, existing staff can be used. It is possible for a single researcher to collect critical incidents, even for multiple research sites (see Ozkaramanli, 2005; Radford, 1993, 1996, 1999). For a project with a larger, perhaps national, scope, it is necessary to gather a team of researchers (Redman, Lambrecht, & Stitt-Gohdes, 2000).

It is important to develop a plan for recruiting subjects. In the public or academic library setting it may be necessary to over-recruit and to offer an incentive such as the $15.00 value card issued to student interview participants reported by Jacoby and O'Brien (2005). Even with this incentive, only five of twelve students recruited actually showed up for the interview (Jacoby & O'Brien, 2005). Group interviewees are frequently recruited with offers of refreshments. Budget considerations usually dictate the range of incentives that can be offered.

Stage 3: Collecting the Data

For all data collection methods, a data collection instrument (survey or interview schedule) must be constructed and training materials or instructions for the data collection team must be developed. It is critically important that those collecting the data be thoroughly trained. Redman, Lambrecht, and Stitt-Gohdes discuss specifics of conducting the CIT interview, which they recommend as a "powerful tool" for data collection (2000, p. 136). For group or individual interviews, it is highly recommend that they be audio-taped with permission of the participants (see Radford, 1993, 1999).

In determining what questions to ask, it is useful to visualize the final

report or paper to be written and to think about what questions will help to gather the information needed for that report. In commercial marketing, CIT questions could be as simple as "What did you like most (least) about your airline flight today? Why?" For the CLASP project, each student was asked to recall and describe in their own words: (a) a successful library experience either recently or in the past (that is, a positive critical incident), (b) an unsuccessful library experience (that is, a negative critical incident), and (c) the factors that made the experience successful or unsuccessful. Flanagan (1954) reported that participants provide 10 percent more incidents if asked about positive incidents first. Additionally, appropriate demographic questions should be included in the survey instrument to help in describing the subjects. The CLASP questionnaire included demographic questions as to the participant's gender, grade level, and languages spoken at home (see Appendix A for CLASP questionnaire).

Decisions on when to collect the data and how much data to collect must also be made. "Data collection may take place while the activity is ongoing, e.g., by supervisors, or by reports of fairly recent activity" (Fisher & Oulton, 1999, p. 114). In the case of CLASP, the data was collected after a series of school visits, gathering perceptions and recollections of the student subjects. Because of the nature of CIT research, large numbers of incidents are frequently collected, but this varies widely. For example, Hamer (2003) collected and analyzed eight CIT interviews of approximately seventy-five minutes each. On the other end of the spectrum, research reported here analyzed 2,416 critical incidents for the CLASP program evaluation. In deciding on sample size it is important to realize that the amount of time required to analyze large amounts of qualitative data whether collected by questionnaires or interviews requires significant commitment of budgetary and human resources.

Flanagan (1954) provided some guidance on sample size, indicating that it must be determined based on the type of activity to be studied. "The underlying rationale is not to be able to make statistical generalizations but rather to ensure that the whole content domain is covered" (Fisher & Oulton, 1999, p. 114). Decisions about how many incidents to collect are unique to each project and depend on such factors as available budgetary and staff resources, purpose and intended use of critical incident results, and target audience for findings. One strategy is to decide upon a range (for example, 50 to 100) of incidents, to collect the minimum number, to begin analysis, and to see if content categories are saturated. If new categories are continuing to emerge with each incident, continue data collection. If not, stop.

The CLASP sample was drawn from all participating schools (private and public) throughout New York City. It was selected in consultation with representatives from The New York Public Library, the Brooklyn Public Library, and the Queens Borough Public Library and took into consideration

which schools would be willing to participate and facilitate distribution of the questionnaires. A minimum of one class of fifth and/or seventh graders was chosen from each of the 68 selected schools, thus determining the resulting number of 2,416 students in the sample. All CLASP staff who assisted in distribution of the questionnaires were given written instructions (see Appendix B for CLASP instructions) and, in March of 2000, all questionnaires were administered.

Stage 4: Analyzing the Data

By far the stage that requires the largest investment in time is stage four. The goal of analysis for any research project is to make sense of a large mass of data through data reduction techniques that summarize and describe the data efficiently (Fisher & Oulton, 1999). For qualitative data, content analysis is frequently performed to identify common themes within the data (see Hamer, 2003; Radford, 1993, 1999).

During stage four the data is transcribed (if interviews were audio-taped) or typed into a word processor or software package. Data should be organized into files or notebooks with each critical incident being coded with a unique number. Incidents are then carefully read and sorted into content themes in an iterative process. Much has been written on the process of analyzing qualitative data (for example, see Strauss & Corbin, 1998). Content analysis and theme development is a subjective process, requiring that more than one researcher be involved in data analysis. Researchers generally use a second or third rater to look at theme development for a portion of the data (10 percent to 20 percent or more). Having a high level of inter-rater agreement adds reliability to the findings (see Radford, 1999; Redman, Lambrecht, & Stitt-Gohdes, 2000).

One way to organize the analysis is to sort the data by research question. In the CLASP study, the theoretical foundation of Watzlawick, Beavin, and Jackson (1967) provided a framework to begin sorting the data by "relational" or "content" dimensions (see Radford 1993, 1996, 1999 for a description of this process of developing themes using the Watzlawick, Beavin & Jackson framework). Software packages are available to assist in this process. [3] Data can also be analyzed by hand or by using a spreadsheet (see Redman, Lambrecht, & Stitt-Gohdes [2000] for discussion of use of a code sheet for hand or spreadsheet analysis). One technique for manual sorting of large amounts of textual data is the Multichromatic Analysis Technique (MAT), which "involves use of colored markers and colored paper clips as aids to analysis" for large qualitative data sets (Radford, 1999, pp. 46–47; see Radford, 1993 or Skiba-King, 1993 for a description of the MAT developed by de Chesnay).

Using the MAT, CLASP student responses to the critical incident questions were sorted into categories. Responses to the survey questions to provide "good" and "bad or unpleasant" experiences were sorted into large

preliminary categories. Responses were further grouped into nine content (information) and four relational (affect) themes following the work of Watzlawick, Beavin, and Jackson (1967), Goffman (1959) and Radford (1993, 1996, 1999). Content-oriented statements tended to focus on information exchange while relational-oriented statements dealt with feelings and attitudes of the participants, which were indicative of the type of relationship between the librarian/staff member and preadolescent user.

Stage 5: Interpreting and Reporting

It is important to document method and decision points along the way. If this is done in a consistent manner, the report will be easier to write. The audience should be kept in mind when constructing the research or project report. A two-page executive summary should be provided as an introduction to a report written for administrators or funding agencies. When writing for publication, detailed method and procedures descriptions are extremely important. For all audiences, the results and interpretation (discussion section) will be of high interest, so emphasize these. For interpretation of results, a conceptual framework is recommended. "Interpretation of the findings is dependent upon a solid grounding in a conceptual base that allows conclusions and educational implications to be drawn from the thematic base" (Redman, Lambrecht, & Stitt-Gohdes, 2000). A return to the stage one review of the appropriate literature, or an expanded review, may be fruitful in suggesting a framework for interpreting data.

The findings should be presented in a form that is useful to the intended audience and should be accompanied by an evaluation of the limitations, validity, and reliability of the method (Fisher & Oulton, 1999). When reporting results, numerical counts of theme frequency can be provided, or alternatively, themes can be listed in order of frequency but without the actual counts. Representative quotations from participants for each theme should be identified and included in the report to provide interesting illustrations and to help the reader understand and contextualize the findings. The report below of the CLASP results gives the numerical counts of theme frequency and includes representative quotations.

Results of CLASP CIT Analysis

A total of 2,452 questionnaires were distributed and returned for a return rate of 100 percent. Of these, 36 were judged unusable (illegible or all questions left blank). There were 2,416 usable questionnaires for analysis (98 percent),[4] an unusually high number for survey research. The high return rate and number of usable questionnaires is thought to be due to the method of distribution in which the CLASP librarians handed out and collected the questionnaires in each selected class. Overall, students were remarkably forthcoming, and their written responses represent their feelings and beliefs quite eloquently, as seen in their quotations below.

Questionnaires were received from a citywide total of 68 schools in 20 districts. There were 382 usable questionnaires from Manhattan, 557 from the Bronx, 93 from Staten Island, 218 from Brooklyn, and 1,166 from Queens. Of the usable questionnaires, 1,077 (45 percent) were from boys and 1,291 (53 percent) were from girls; 48 (2 percent) did not indicate gender. There were 1,270 (53 percent) fifth grade students and 1,146 (47 percent) seventh grade students. The students were from diverse ethnic backgrounds, seen in the variety of languages spoken at home. English alone was spoken in 428 homes, Spanish alone in 68 homes, and both Spanish and English in 309 homes. A total of 60 additional languages and language combinations were spoken in 1,591 homes.

Positive Critical Incidents—Content Themes

What do students remember as being critical to their having had a successful visit to the library? A total of 1,680 students responded to the positive CIT question that asked them to think about times they had a good experience in the public library and to write down what happened and what they thought it was that made this visit good. Students reported that they had positive experiences in the library when they found the books or information they were looking for, used the computers, attended library programs, and the library atmosphere was pleasant. The content- (information) related responses were grouped into nine themes: Found Good Book, Used Computer/Internet, Found Information for School Assignment, Attended Library Programs, Got Library Card, Enjoyed Atmosphere/Facilities, Found Good Videos, Learned Information/Skills, and Found Different Formats. These are listed below, with representative student quotes as examples.[5]

Found Good Book(s) (468)
When I went to the library after school. I was doing my homework. I saw a good book. I was so delighted, I borrowed it. This was a good visit to the library because I felt like I was inside the book. (fifth grade boy)

Used Computer/Internet (327)
Being able to go on the internet. Don't have computer at home so I'm not able to go on the internet very often. (seventh grade boy)

Found Information for School Assignment (for example, Reports, Projects) (250)
I needed to find information so I could finish my report and they had everything I needed. I can always count on the library to get the right information for me. (seventh grade girl)

Attended Library Programs (133)
One time when I was little I went to the library and I sat down and listened to a story where the person who said it did origami at the same time. It was a good story and me and my brother liked it a lot. (fifth grade girl)

Got Library Card (98)
When I got my library card and I found out I could take out up to 10 books. (seventh grade girl)

Enjoyed Atmosphere/Facilities (74)
When I first went to the library I loved it because it was quiet and peaceful and the librarians let me go on the computers and I made an essay and report on black history. (fifth grade, gender not indicated)

Found Good Video(s) (54)
[It was a good experience in the library] when I got my adult card. Also when I started to take out videos. I had never done this before, so it was very exciting. (seventh grade boy)

Learned New Information/Skills (49)
One time I needed an article from the *New York Times* and they gave me a roll of a newspaper on a film. It made this a good visit when they taught me how to use it on a machine. (seventh grade girl)

Found Different Formats (16)
It was the first time I went to a library. I was so happy because I found a lot of cool books, CDs, movies, and computers. I got a 100 on this small project and it was fun doing my work in the library. (seventh grade boy)

Positive Critical Incidents—Relational Themes

Students also reported remembering good visits when they were treated well by librarians, felt good about themselves, enjoyed the company of friends and family, and had a variety of positive experiences. These relational responses were sorted into four relational (affect) themes: Librarian Attitude, Social Aspects, Positive Experiences/Emotions, and Felt Good about Self. They are illustrated with representative examples below.

Librarian/Staff Attitude (275)
When a guy in the [branch] library helped me find information on my research paper. Even though there were a lot of people he helped everyone. He even helped me look at the passage to see if there was any information I needed. It made a good visit because [before] when I went to the library to ask for help they just wrote a number and [said to] look it up for yourself. (seventh grade girl)

Social Aspects (188)
Once I went to the library, and I met my friend there. We researched for our project. We asked the librarian and she helped us. [What made it a good visit was] that I went to the library with my friend & we both helped each other. After that we went out to eat. (seventh grade girl)

Positive Experiences/Emotions (108)
When I was picking out a book on money I found real money in the book. [It was good] because if I didn't go to the library I would've never found the money. (fifth grade boy)

Felt Good about Self (50)
Someone was looking for a book and they didn't know where it was so I said "you need some help" and I got the book for him. I felt [like a] grown woman and I felt like a librarian. (fifth grade girl)

Negative Critical Incidents—Content Themes

What do students remember as being critical in determining unpleasant experiences in the library? In order to find the answer to this question, students were asked to think about times they had visited the public library and to remember a time when they had an unpleasant experience in the library. They were asked to write down what happened and what they thought it was that made this an unpleasant visit. Nine hundred and eighty-four student responses to this question were analyzed and sorted into five content (information) and four relational (affect, emotional) themes. It is important to note that 17 percent of the 984 students responded that they never had a bad experience in the library.

Students reported that they had unpleasant experiences in the library when they could not find the books or information they were looking for; had problems at checkout (usually related to their having overdue books or fines); had problems using or getting access to the computers; or the library atmosphere was problematic (usually noisy or crowded). The content- (information) related responses were grouped into five themes—Could Not Find Information, Checkout/Procedural Problems, Computer Problems, Library Atmosphere Problematic, and Negative Experiences—and are illustrated with representative examples below.

Could Not Find Information (150)
A time it had not been pleasant is when I went to the library to look for a book and it is not there. I do not feel good when that happens. What made this a bad visit to the library was they didn't have the book that I needed and I felt that all libraries should have everything you need. (fifth grade girl)

Checkout/Procedural Problems (94)
I remember now it was when they looked at my file and found out I had a list of overdue books and I had to pay a big fine. They found out that I had a list of overdue books waiting to be paid. At first I was nervous and scared because I thought they would just get mad at me and make me pay right away. (fifth grade boy)

I took out a book and they told me that there were some books I didn't return. The lady at the front was rude. She kept on telling me that I owed stuff I never took out. (seventh grade girl)

Computer Problems (68)
When I went to the library I couldn't find any books on my project. So, I signed on the computer, but I had to wait around 2 hours. At home my printer was broken, and I had to wait 2 hours. The librarian was mean too. (seventh grade girl)

Library Atmosphere Problematic (47)
Once when I went to the library everybody was talking loud and was yelling. The visit was unpleasant because I went there to study and everybody was talking so I couldn't concentrate. (seventh grade boy)

Negative Experiences (30)
The printer took my money. (fifth grade girl)
My parents forced me to go. I couldn't play that day. I was missing all the fun. I could have been playing. (seventh grade boy)

Negative Critical Incidents—Relational Themes

Students reported that they had negative or unpleasant experiences in the library related to how they were treated by the librarians/staff; when they encountered other users who were problematic; when they got in trouble, were reprimanded, or were asked to leave ("kicked out"); or when they were embarrassed or afraid. The relational responses were grouped into six themes: Librarian/Staff Attitude Poor, I Got "Kicked Out," Other Users Problematic, I Got in Trouble, Negative Emotions, and I Was Embarrassed; they are illustrated with representative examples below.

Librarian/Staff Attitude (152)
Oh yeah, the first time I went there was a very mean lady there. She was very bad and for nothing yelled at everyone and on purpose. She kicked 5 people out. And I never went back to my library. (seventh grade girl)
The librarian kicked me and my friends out, but we weren't the ones talking. The librarian was so rude and was accusing the innocent. (seventh grade girl)

I Got "Kicked Out" (71)
When I was talking then they said "be quiet" and I was quiet, but they kicked me out anyway. [It was a bad or unpleasant visit because] I was not talking but they kicked me out anyway. (fifth grade boy)

Other Users Problematic (59)
A bad experience was when some teenagers came to me and were very rude and obnoxious. The teenagers were out of control and the librarians didn't stop or say anything to them. (seventh grade girl)

I Was Embarrassed (42)
I felt embarrassed when I tripped and fell and everybody saw me. (seventh grade boy)

I Got in Trouble (25)
I got in trouble because I stayed on the computer 5 minutes too long. The librarian kept on screaming because I was eating gum. (fifth grade girl)

Negative Feelings (19)
A bad time in the library was when I could not read. I just got a book and was just looking at the pictures. I did not know how to read or spell. (fifth grade boy)

The CIT: Discussion and Limitations

The above report of results with representative quotations shows the richness of the results from the CIT method and how it can be used for program evaluation and to formulate specific recommendations for program improvement. Based on the CIT evaluation of CLASP, numerous recommendations were developed to enhance and foster positive student experiences in the library and to assist students in having fewer negative experiences.

One of the most interesting findings from this research is that the technique of having fifth and seventh grade urban students write critical incidents is such a productive method of data collection.[6] Since the population is so diverse, with so many languages represented, one might expect that the students would not be as articulate as is demonstrated here. In general, fifth graders were found to be more forthcoming in their answers and wrote longer answers. This finding may be due to developmental differences in which the younger students are not as reluctant as the older ones to reveal their emotions. Interestingly, boys were just as likely as girls to be forthcoming and to give answers related to relational as well as content topics.

Despite some reluctance, which is to be expected in preadolescent populations, students revealed themselves to be quite sensitive. They can be easily embarrassed, have a keen sense of fairness, and resent injustice, especially when they feel that they have been falsely accused. They report vivid memories of times when they were "yelled at," "scolded," or asked to leave and write that they were, at times, deeply affected by their fears and insecurities. For example, one seventh grade girl who was asked to leave the library during her first visit wrote: "And I never went back to my library." She felt that she had been unfairly reprimanded by a librarian or staff member, so she left and chose never to return (see also Radford & Radford, 1997).

Another noteworthy finding is that for both positive and negative critical incidents, the largest numbers of responses for the relational category were

centered on "Librarian/Staff Attitude." Thus, for these preadolescents, the most important factor in their perception of successful library visits is the attitude of the librarian or staff member they encounter. They appreciate it when the librarian/staff member takes an interest in them, shows flexibility in bending the rules or makes exceptions for them (for example, extends computer time for their assignments), and accompanies them to find a book on a shelf instead of pointing or providing a call number on a slip of paper. Radford (1993, 1996, 1999) reported similar findings regarding college-age academic library users. Thomas notes that "in some cultures the relational 'work' to establish a basis for further interaction must precede information tasks" (1999, p. 159), which is affirmed by the data analysis above. Students need to be reassured that they will have a positive interaction before they feel comfortable enough to approach the librarian/staff person or reveal their information need (see also Kuhlthau, 2004). Furthermore, Bialo and Sivin-Kachala (1996) assert that relational, interpersonal aspects of library encounters are especially important for school librarians serving culturally and ethnically diverse school populations.

The application of communication theory to interpersonal encounters of preadolescents in the library context is at the exploratory stage. One criticism of library literature has been its lack of robust theory. The CIT provides a method that enables the application of the heuristically rich relational theory from the communication field to library interactions. In addition, collecting critical incidents elicits the user's point of view to advance knowledge of the preadolescent perspective in library interactions.

The results of this CIT analysis argue for a new model of the librarian-preadolescent reference interaction that takes a process approach and recognizes the vital importance of the interpersonal, relational messages that are communicated in the encounter along with the transfer of information, instructions, or suggestions for research strategies (see also Kuhlthau, 2004). It also integrates the user's perspective as critical to understanding this complex encounter (Morris, 1994). Related research in virtual reference (chat) encounters (such as Ask-A-Librarian services) has found that interpersonal aspects similar to those of face-to-face interactions are present, and they are extremely important to success, especially in interactions with adolescents and preadolescents (see Radford, 2006a, 2006b; Radford & Thompson, 2004).

Regarding practical implications, this study suggests that education for children's and young adult librarians could benefit from increased study of interpersonal dynamics. According to Winston and Paone:

> Young adults currently represent a large and important segment of the user population in public libraries and population estimates indicate that this group will be a growing segment as well. However, the research literature and practice-oriented literature in library and information science have not addressed a number of issues associated

> with the service provision and characteristics of this user population. (2001, p. 49)

If librarians were made more aware of the impact of their actions on the fragile egos of preadolescents, they might be more cautious in administering discipline, striving not to "accuse the innocent," or inadvertently driving students away from the library or discouraging them from asking questions. Furthermore, cultivating positive relationships with preadolescents and adolescents will multiply the pleasant interactions and minimize the negative ones. Winston and Paone (2001) found that public libraries did not give priority to services to the young adult population and that it is necessary to maintain sufficient staffing levels, especially age-level specialists, to meet their needs. Whether or not public service librarians are age-level specialists, it is possible to improve the quality of encounters with youthful clients. Although many library practitioners believe that "people skills" are inherent, research findings demonstrate that librarians can be educated to improve their interpersonal skills in reference encounters (Dewdney, 1987).

This research is exploratory, and, although a large number of students were surveyed, no claim for generalization of these results is made. Limitations of the CIT research design include the fact that data are self-reported, in this case by students, and, as such, are subjective accounts of their perceptions. In addition, the schools surveyed were not selected randomly but were chosen in recognition of access and facilitation issues. Surveys were administered and collected by CLASP librarians, which may have impeded some students from answering candidly.

LIS Research Using the CIT

According to Fisher and Oulton, "The Critical Incident Technique has been tried and tested in a wide range of discipline areas and for a variety of purposes. It is recognized as a valid, reliable and effective method for gathering rich qualitative data for a variety of purposes" (1999, p. 126). The CIT has been successfully used in many social science fields, in marketing and business applications, and in the LIS context (see Andersson & Nilsson, 1964; Andrews, 1991; Carr, 1980; Fivars, 1980; Radford, 1999; Shirey, 1991). A brief overview of some of the projects in which the CIT is used in LIS is provided here to illustrate the variety of these applications.

Fisher and Oulton's (1999) article provides an excellent overview of the CIT and also describes three studies by researchers in the UK that apply the technique to research: staff development needs in the context of change in UK higher education, decision-making practice in European libraries, and developing a tool to support library workers entering management positions.

Hamer (2003) used the CIT to investigate information seeking of young gay males regarding coming out and took a social constructionist perspec-

tive of gay identity. CITs were collected in audio-taped interviews lasting approximately seventy-five minutes with eight volunteers in their late teens or early twenties. Results indicate that subjects had three types of information needs: self-labeling, consequences for self-identifying as gay, and forming an understanding of gay identity. In addition, Hamer found that their information seeking was characterized by the experience of fear and by the need for concealment and secrecy.

In her doctoral dissertation, Ozkaramanli (2005) used the CIT to study the perceptions of librarians of quality digital reference services. Critical incidents were collected through interviews with forty librarians from ten academic libraries in Ohio and Pennsylvania that offered chat reference services. Findings revealed that librarian and user attitudes, question negotiation, and availability of resources were critical to perceived success in chat interactions. Ozkaramanli (2005) provides a detailed explication of the CIT and her data analysis technique.

Radford (1993, 1996, 1999) used the CIT to study interpersonal communication aspects of reference service in academic libraries, collecting forty-seven critical incidents from twenty-seven academic librarians and twenty-seven library users. Radford's work was based on the theoretical foundation of Watzlawick, Beavin, and Jackson (1967) and Goffman (1959) that differentiated the dual nature of messages as having both content and relational dimensions. Watzlawick, Beavin, and Jackson's (1967) perspective draws attention to the idea that more than correct answers to reference questions are being communicated in reference encounters. In addition to providing this content information in response to a user's request, librarians are also communicating relational information in their verbal and nonverbal expressions that has additional meaning for library users.

Conclusion

The CIT is a flexible tool able to be applied in a variety of settings and for a variety of purposes. As can be seen in the above results and discussion, the CIT has provided a method for tapping into student's perceptions, yielding a substantial data set for program evaluation and a qualitative analysis of the communication process. Qualitative measures, such as the CIT, help to capture the differences that may fall between points on a standard scale. They can answer such questions as "What do programs mean to participants? What is the quality of their experience?" (Patton, 1987, p. 30). They capture nuances of quality that are lost in most survey data collection. Clearly, the CLASP study has only begun to explore the dynamics of the complex interaction between librarians, library staff, and young people. With increased understanding of this process, greater success and satisfaction for both preadolescent users and librarians can be achieved. The CIT provides a method for expanding one's understanding of this interaction, especially from the young person's point of view. It is hoped

that this article will stimulate interest in the CIT and will encourage LIS researchers to consider adopting this method in future projects, including program evaluation.

Notes

The author would like to thank The New York Public Library, Brooklyn Public Library, and Queens Borough Public Library for permission to publish these findings. The author also extends heartfelt thanks to Kate Todd, Grace Shanrahan, and Margaret Tice for their guidance and assistance and to all the CLASP librarians and students who took part in the evaluation project.

1. The CLASP evaluation included survey questionnaires about the CLASP program, including questions designed to document the impact of the CLASP visits on student's perceptions of the public library. Other than the CIT questions, these results are not reported here in detail (Radford, 2000; Tice, 2001).
2. The CLASP evaluation included survey questionnaires about the CLASP classroom visits and other aspects of the program. As noted above, other than the CIT questions, these results are not reported here in detail. However, it should be noted that the results were very positive, with 75 percent of students surveyed remembering CLASP visits and large numbers of students reporting positive impacts of CLASP (Radford, 2000; Tice, 2001).
3. There are software packages available to assist in analysis, such as NVivo, MAXqda, and Atlas.ti, but such packages require a substantial investment of funds and time to learn how to use them effectively. See Miles and Huberman (1994) for a description of this type of software product.
4. All percentages are rounded to the nearest whole number.
5. For all quotations throughout this paper, the number of responses is given in parenthesis and themes are listed in descending order of frequency. Note that student answers can be sorted into more than one theme when more than one concept is present. Students' minor errors in spelling or grammar have been corrected and words have been added [in brackets] as necessary for clarification.
6. It is to be noted that research involving preadolescents' and adolescents' perceptions of librarians has been scarce, although some researchers have investigated this area from a qualitative perspective in the school library context (for example, Chelton, 1999, 1997; Mellon, 1995) and through survey research in public libraries (for example, Winston & Paone, 2001).

References

Andersson, B., & Nilsson, S. (1964). Studies in the reliability and validity of the critical incident technique. *Journal of Applied Psychology, 48,* 398–403.

Andrews, J. (1991). The use of the critical incident research technique in an academic library. *Library & Information Research News, 14*(50), 22–27.

Arnold, M. J., Reynolds, K. E., & Ponder, N. (2005). Customer delight in a retail context: Investigating delightful and terrible shopping experiences. *Journal of Business Research, 58*(8), 1132–1145.

Bialo, E. R., & Sivin-Kachala, J. (1996). The effectiveness of technology in schools: A summary of recent research. *School Library Media Quarterly, 25*(Fall), 51–57.

Bycio, P., & Allen, J. S. (2004). A critical incidents approach to outcomes assessment. *Journal of Education for Business, 80*(2), 86–92.

Carr, D. (1980). The agent and the learner: Interactions in assisted adult library learning. *Public Library Quarterly, 2*(2), 3–19.

Chelton, M. K. (1997). The "overdue kid:" A face-to-face library service encounter as ritual interaction. *Library & Information Science Research, 19*(4), 387–399.

Chelton, M. K. (1999). Structural and theoretical constraints on reference service in a high school library media center. *RUSQ, 38*(3), 275–282.

Dewdney, P. (1987). *User satisfaction with the public library reference interview: An experiment to determine the effects of training librarians in communication skills.* Unpublished doctoral dissertation, School of Library and Information Service, University of Western Ontario.

Fisher, S., & Oulton, T. (1999). The critical incident technique in library and information management research. *Education for Information, 17*(2), 113–126.
Fivars, G. (1980). *The critical incident technique: A bibliography* (2nd ed.). Palo Alto, CA: American Institutes for Research in the Behavioral Sciences. (ERIC Document Reproduction Service No. ED 195 681).
Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 5,* 327–358.
Goffman, E. (1959). *The presentation of self in everyday life.* Garden City, NY: Doubleday Anchor.
Hamer, J. S. (2003). Coming-out: Gay males' information seeking. *School Libraries Worldwide, 9*(2), 73–79.
Jacoby, J., & O'Brien, N. P. (2005). Assessing the impact of reference services provided to undergraduate students. *College & Research Libraries, 66*(4), 324–340.
Kuhlthau, C. C. (2004). *Seeking meaning: A process approach to library and information services* (2nd ed.) Westport, CN: Libraries Unlimited.
Mellon, C. A. (1995). "She's nice and she helps:" What 6th graders say about school librarians. *School Library Journal, 41*(5), 27–29.
Miles, M. B., & Huberman, A. M. (1994). *An expanded sourcebook: Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage.
Morris, R. (1994). Toward a user-centered information service. *Journal of the American Society for Information Science, 45*(1), 20–30.
Ozkaramanli, E. (2005). *Librarians' perceptions of quality digital reference services by means of critical incidents.* Unpublished doctoral dissertation, University of Pittsburgh.
Patton, M. Q. (1987). *How to use qualitative methods in evaluation.* Newbury Park, CA: Sage.
Radford, M. L. (1993). *Relational aspects of reference interactions: A qualitative investigation of the perceptions of users and librarians in the academic library.* Unpublished doctoral dissertation, Rutgers, The State University of New Jersey.
Radford, M. L. (1996). Communication theory applied to the reference encounter: An analysis of critical incidents. *Library Quarterly, 66*(2), 123–137.
Radford, M. L. (1999). *The reference encounter: Interpersonal communication in the academic library.* Chicago: ACRL, a Division of the American Library Association.
Radford, M. L. (2000). *CLASP—Connecting Libraries and Schools Project, 1999–2000.* Unpublished report for The New York Public Library, Brooklyn Public Library, and Queens Borough Public Library, May 23.
Radford, M. L. (2006a). Encountering virtual users: A qualitative investigation of interpersonal communication in chat reference. *Journal of the American Society for Information Science and Technology, 57*(8), 1046–1059.
Radford, M. L. (2006b). Interpersonal communication in chat reference: Encounters with rude and impatient users. In R. D. Lankes, E. Abels, M. White, & S. N. Haque, (Eds.), *The virtual reference desk: Creating a reference future* (pp. 41–73). New York: Neal-Schuman.
Radford, M. L., & Radford, G. P. (1997). Power, knowledge, and fear: Feminism, Foucault and the stereotype of the female librarian. *Library Quarterly, 67*(3), 250–266.
Radford, M. L., & Thompson, J. (2004). "*Yo Dude! Y R U Typin So Slow?"* Online proceedings of the Virtual Reference Desk 6th Annual Conference, Cincinnati, OH, November 8–9. Retrieved May 21, 2006, from http://www.webjunction.org/do/DisplayContent?id=12499.
Redman, D. H., Lambrecht, J. J., & Stitt-Gohdes, W. L. (2000). The critical incident technique: A tool for qualitative research. *Delta Pi Epsilon Journal, 42*(3), 132–53.
Ruben, B. D. (1993). What patients remember: A content analysis of critical incidents in health care. *Health Communication, 5,* 1–16.
Shirey, D. (1991). Critical incident technique. In A. Kent, (Ed.), *Encyclopedia of Library and Information Science* (Vol. 6, pp. 286–291). New York: Dekker.
Skiba-King, E. (1993). *An examination of the patterns of self-reported disclosure by incest survivors.* Unpublished doctoral dissertation, Rutgers, The State University of New Jersey.
Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.
Thomas, N. P. (1999). *Information literacy and information skills instruction: Applying research to practice in the school library media center.* Englewood, CO: Libraries Unlimited.
Tice, M. (2001). Queens Borough Public Library and the Connecting Libraries and Schools Project. *Journal of Youth Services in Libraries, 14*(3), 11–13.

Watzlawick, P., Beavin, J., & Jackson, D. (1967). *Pragmatics of human communication*. New York: Norton.

Winston, M., & Paone, K. L. (2001). Reference and information services for young adults: A research study of public libraries in New Jersey. *Reference & User Services Quarterly, 41*(1): 45–50.

Marie L. Radford is an Associate Professor at Rutgers SCILS. Prior to joining the Rutgers faculty, she was Acting Dean and Associate Professor of Pratt Institute's School of Information and Library Science. She holds a Ph.D. from Rutgers University and an M.S.L.S. from Syracuse University. Before arriving at Pratt in 1996, she was Head of Curriculum Materials at William Paterson University of New Jersey's Cheng Library and a school librarian in Hunterdon County, NJ. Her research interests are interpersonal communication aspects of reference service (traditional and virtual), evaluation of digital resources and services, qualitative methodologies, cultural studies, and media stereotypes of librarians. She has given numerous international, national, and local presentations at scholarly library and communication conferences. She has published in *Library Quarterly, College & Research Libraries, Library Trends,* the *Journal of Education for Library & Information Science, JASIST,* and the *Journal of Academic Librarianship.* Her latest book, *Web Research: Selection, Evaluation, and Citing* (2nd ed.) was published by Allyn and Bacon in 2006. Her book *The Reference Encounter: Interpersonal Communication in the Academic Library* was published by ACRL/ALA in 1999. She is active in library and communication associations, including the ALA, NJLA, ALISE, and RUSA. She is the past chair of the Library Research Round Table and presently a member of the LRRT nominating committee. She is presently a member of the RUSA Publications Committee and the ALA Research & Statistics Committee. She co-chaired the ALISE conference program in 2004 and is a member of the program committee for the upcoming Library Research Seminar IV.

APPENDIX A CLASP QUESTIONNAIRE
Connecting Libraries and Schools Project (CLASP)
STUDENT QUESTIONNAIRE

School:_____
District:_____
Borough:_____
Class:_____

DIRECTIONS: Please write your answers to the questions below. Try to answer all questions. Use your best handwriting or printing. Do not put your name on this form; all answers will be kept private. Your answers are very important to us. Thanks!

Gender (circle one): Girl Boy
Grade Level (circle one): 5th 7th
Languages (circle all languages spoken at home):
English Spanish Chinese Russian Creole Urdu Farsi Other_____

1. Think about times you have visited the public library. Remember a time when you had a good experience in the library. Please write down what happened.

_____
_____

1A. What was it that made this a good visit to the library?

_____
_____

2. Now remember a time when you had an unpleasant or bad experience in the library. Please write down what happened.

_____
_____

2A. What was it that made this a bad or unpleasant visit to the library?

_____
_____

3. Can you remember a time when a public librarian came to your classroom to talk about the library? Circle one: YES NO

3A. If yes, have the public librarian visits changed your experiences in the public library? Circle one: YES NO

3B. If yes, how have the librarian visits changed your experiences in the public library?

_____

_____

4. Is there anything else you would like to tell us about the public library?

_____

_____

Thank you for helping us to make your public library better for you and for other students.

APPENDIX B CLASP INSTRUCTIONS

Instructions for CLASP Staff
Student Questionnaire Collection
Here are instructions for the person administering the questionnaire:

1. Give 1 copy of the questionnaire to each student.
2. Tell the students: "This is not a test. There are no right or wrong answers."
3. Tell the students: "If you answer these questions, it will help us make the library better for you."
4. Tell the students: "Do not put your name on the paper. Your answers will be private and confidential."
5. Read each question to the students, starting with School. Make sure the students answer every question. If necessary, help them to understand what the question means. Give them about 4 or 5 minutes for each essay answer.
6. Collect all forms at the end of the class period. Send the forms immediately to EP/CLASP. Do not wait until you have visited all the schools. Complete the collection of all forms before March 31, 2000.

Discourse Analysis and the Study of Communication in LIS

John M. Budd

Abstract

A considerable portion of the work that is done in library and information science (LIS) can benefit from discourse analysis as a research method. The two major families of discourse analysis are linguistic-based analysis (such as conversation, which could be applied in any setting where information professionals mediate between the universe of information and information seekers), and culturally or socially based discursive practices (along the lines of the analyses that Michel Foucault has conducted). The potential of both families for LIS inquiry, along with examples of both, are discussed.

Two Varieties of Discourse Analysis

Version 1. I want you to believe me. For you to believe me, I have to be credible to you. To be credible to you, I must speak in lexical terms that are familiar to you; I have to be understood. The lexical comprehension is one part of understanding; I also have to communicate in ways that fit your knowledge base, that will have a context within your mind. Understanding is one path to belief; it is necessary, but it is not sufficient. Your belief in what I say is also dependent upon your acceptance of what I say. I have to persuade you that what I say is correct; I must employ rhetoric as a means of setting you on a path of agreement that will culminate in your belief in what I say. At any point in our exchange you might reject what I say; you might disagree, perhaps strongly enough that you immunize yourself against all rhetorical strategies and tactics I can muster. In short, your belief in what I say may be hard-won, may be given, may be tenuous, may be impossible. Your belief in what I say is based in a complex array of discursive events—what we say to one another, what has been said to you

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 65–82
© 2006 The Board of Trustees, University of Illinois

in the past (directly, as in conversation or presentation, or obliquely, as in your reading of previously written texts), what you have thought, and what you have said.

The foregoing can describe, among other things, a reference interview in a library. If you are a student or a community member asking a question, the above conditions tend to apply in a discursive exchange. Further, the facets of the exchange can be examined rigorously so as to fix the locus of success or failure in such an exchange.

Version 2. I want you to believe me. I still have to be credible to you. In order to accomplish this credibility I will call upon traditions, customs, sources, powerful institutions, and other necessary social relations. I will ensure that you believe me by making it impossible for you to disbelieve me. What I say will build upon a substantial accumulation of discourse that has been established as authoritative. You believe me because you believe that set of discursive practices. The practices are not a continuous line from the past, although they have roots in the past. Their history has been disjointed, but it has managed to gain acceptance over, and through, time. All of your affiliations influence your belief structure—your education, your political party, your geographic location, your religious views, your occupation, your family, your friends, and your economic status.

This version is no less complex than the first; the main difference is that these discursive practices are not usually individual, dyadic, or engaged in by small groups. The practices in the second version are usually formal, whether written or spoken. They tend to be actions in the forms of making speeches, writing articles and books, issuing proclamations, and publishing results of inquiry. All disciplines engage in these practices, including library and information science (LIS). It is also possible to examine our own discourse rigorously and according to exemplars of analysis.

A few things must be mentioned at the outset of this article. The first is that, while there are many ways to study discourse and many purposes of each study, the focus here will be on two families of discourse analysis. The first is the more traditional, linguistic-oriented examination that can be framed as a form of applied semantics. Conversational analysis is an example of this type. The thrust, simply put, is investigation into what people say as part of efforts to be understood by, and to understand, others. The second family attends more to social, political, and other aspects of communicative practice. The aim is frequently to place discourse within a context or milieu, seeking to explicate not merely surface meanings of statements but possible structures into which utterances may fit. For both of these families, discourse is language beyond the clause or sentence level; discourse is a larger linguistic unit (Stubbs, 1983). The two families will be examined in some detail, with examples of analyses offered. This article will not be based in an independent analysis of discourse or discursive practice (that would narrow the scope and potential utility severely); it will present ways

of engaging in discourse analysis, reasons why it can be a fruitful method, and what we can learn as a result of it.

Discourse and Language

The first family of discourse analysis (illustrated in Version 1 above) centers principally on what Brown and Yule (1983) call "transactional language." Language used in such a situation is primarily "message-oriented." "In primarily transactional language we assume that what the speaker (or writer) has primarily in mind is the efficient transference of information" (Brown & Yule, 1983, p. 2). Conversations are the most frequently occurring kinds of this language use. For the most part, conversations are structured by both (all) speakers so that they can be taken literally. There may be additional conceits, such as sarcasm or irony, but even those are context bound so that they can be readily understood. Suppose two people are conversing and a portion of their exchange is as follows:

> A:Did you hear what he said?
>
> B:Yeah, but I don't buy it.
>
> A:I don't know; he seemed to know what he's talking about.
>
> B:Yeah, right.

It would be difficult for native English speakers in today's American society to assume that B is actually agreeing with A. At work is what Grice calls "conversational implicature" (1989, p. 26); the conversational context determines the meaning of some words, so "Yeah, right" in the above example is not taken as literal information.

No paper on discourse analysis can ignore the distinction presented by Ferdinand de Saussure (1959) in his *Course in General Linguistics*. He points out that there is a functional difference between *langue* (the linguistic system that provides the structure for any utterance) and *parole* (the real utterances spoken by people within particular situations). Language, says Saussure, "is a self-contained whole and a principle of classification. As soon as we give language first place among the facts of speech, we introduce a natural order into a mass that lends itself to no other classification" (1959, p. 9). Even if we take Saussure's definition of language at face value, its utility in discourse analysis is questionable since the classification he speaks of is less ordered and law driven than he supposes. Saussure's distinction provides a grounding for the applied sociolinguistic analysis of discourse, even when theorists and researchers have disagreed with some fundamental tenets of his theory. The difference between what *could* be said and what *is* said is at the heart of much contextual examination of conversations and other dyadic communication exchanges. Saussure's program, in short, involved attempts to derive formal laws of language based on linguistic structures manifest in speech (which he privileged over writing). He considers speech

(*parole*) to be authentic and writing to be artificial. While his structuralist approach attracts researchers from several disciplines, his fundamental thesis can be found somewhat wanting.

In Saussure's semiotics (or semiology), the sign is the combination of a signifier (a sound-image, or speech sound intended to represent something) and a signified (a concept or thing represented). The sound "dog" signifies the four-legged mammal of the genus *canis*. In French "chien" is the signifier of the same four-legged mammal. For Saussure the sign is arbitrary; nothing, in fact, determines or requires the signifier to be a certain sound or to have an a priori relationship with a thing signified. As John Gumperz explains, "While all information on language ultimately derives from speech, the assumption is that the raw information collected in situ must first be sifted and recoded in more general form before it can be utilized in the linguist's generalizations" (1982, p. 11). Structural linguistics is based on the assumption that speech (*parole*) is to be explained by systems of rules that have functional relationships. For many sociolinguists the foundation of rules of speech may be acceptable but nondetermined relations of signifiers and signifieds may not be.

Semiotics since Saussure presents an even tighter connection between language (in the Saussurean sense of the whole) and discourse. For one thing, critiques of the immutability of language's structure argue that it is *parole* (speech) that should have priority over *langue* (language) in inquiry. Vološinov refutes a central premise of Saussure's theory: "The sign may not be divorced from the concrete forms of social intercourse (seeing that the sign is part of organized social intercourse and cannot exist, as such, outside it, reverting to a mere physical artifact)" (1973, p. 21). The middle ground is perhaps the most effective for a linguistics-based discourse analysis. The sign is not wholly a part of (determined by) the language system; the sign is also not wholly a social construction. It contains elements of both. If it were not *part* of the language system there could be no shared meaning; if it were not *part* of social intercourse there could be no metaphor, simile, metonomy, or—for that matter—poetry or irony. These kinds of speech may be used frequently in exchanges like reference interviews. A librarian might employ similes to enable an information seeker to connect the familiar with the unfamiliar. For example, an undergraduate or high school student may say to a librarian, "I have to write a 500-word paper interpreting what T. S. Eliot might have meant by the line, 'I have measured out my life with coffee spoons'" (Eliot, 1971, p. 5). The librarian could illustrate that the line is not to be taken literally and may refer the student to works that speak to ways of stating certain thoughts, perhaps including someone saying that he or she can mark the last several years of life through the seasons of *Friends*.

The linguistics-based discourse analysis in general draws heavily from a background in the examination of speech acts. J. L. Austin (1975), more

than anyone else, gives the theory of speech acts legs. The theory is elaborated upon by John Searle (1969), whose work will inform this article's look at discourse analysis. We must remember that analysis (of anything) is a formal act that relies on clear and agreed-upon definitions and methods of study. Searle helps provide essential definitions that are integral to discourse analysis. Anyone can appreciate that speech consists of uttering words (or morphemes, to break the unit down even further) that are intended to have meaning. Usually, but not always, the words are strung together in sentences. In events such as reference interviews the speech has another character—it may assert something, or ask a question, or give instructions or commands, or make promises, etc. A single query, such as "Where can I find a biography of Mark Twain?" exhibits this kind of character and is called an *illocutionary act.* A reply such as "You can search the library's online catalog by 'Mark Twain' as a subject" is also an illocutionary act. A reference interview is intended to have some effect; both the questioner and the librarian want the questioner to find a useful biography of Mark Twain. In other words, the set of illocutionary acts that make up the reference interview have an effect on the questioner. The intended effect in this case may be an increase in the questioner's awareness of the details of Mark Twain's life. The increased awareness is called a *perlocutionary act.* These elements, according to Searle, apply whether the utterances are spoken or written, although details of analysis should be sensitive to both means of uttering something.

Applied Analysis and Information

Applied discourse analysis of the type we are focusing on here is, perhaps first and foremost, *not* an idealization of human behavior. Also, it is not a retrospective account by the agents themselves (although such a methodological twist has potential usefulness). It is an examination of actual conversational behavior. As a method in our field, this kind of discourse analysis includes not merely *what* is said but also *how* it is said. The "how" entails the utterances themselves (the words as they are put together in speech), other phonetic sounds that accompany utterances (uh, er, hmm, etc.), and the spaces between utterances. Suppose a teenager approaches a reference desk in a public library and asks, "Do you know where I can find medical books?" The librarian may respond, "Are you looking for something on anatomy, on diagnosis, on diseases and treatments. . . ?" There is a pause of a few seconds and then the teenager says, "Uh, well, I guess I'm looking for books on human reproduction." The pause, plus the "uh" and the "well," carry import and meaning in this kind of exchange. While the foundation of the study of such discourse is linguistic, it would be more correct to say that it is sociolinguistic. The social situation, which both affects, and is reflected in, the exchange, includes the psychological dynamics of the agents. In the above example the librarian will attend to

the pause as well as to the utterance. The analyst will also take note of the pause and include it in the investigation of the exchange of utterances.

The complexity of the analytical element of this discourse analysis necessitates very careful procedural preparation and execution by the analyst. The analyst cannot rely on reports of discursive exchanges by any of the agents involved; their recollections will not reflect either the linguistic or nonlinguistic occurrences with sufficient accuracy. Taping the exchanges will lend a higher degree of fidelity to analysis. A detailed transcript that includes all phonemic and phonetic sounds as well as the timing between the sounds is vital to full examination. Conventions of transcribing exist so that pauses, breathing patterns, simultaneous speaking, and other things can be clearly recorded. The questions, responses, pauses, interruptions, etc., that typify exchanges thus become components of the analysis. Robin Wooffitt observes:

> whereas intuition fails the analyst, recordings of actual events, and detailed transcriptions of them, permits capture of the detail of participants' conduct. The analyst is relieved of the near impossible task of trying to imagine what goes on during the interaction: the analyst can actually find out by careful listening to the tape, and investigation of the subsequent transcript (2001, pp. 50–51).

The discourse, as it exists, is not reducible to abstract linguistic analysis. Tapes and transcripts provide empirical data that can then be interpreted. Discourse analysis can enable a rich and deep examination of how information seekers ask their questions, as well as how librarians answer. One goal of the analysis is the improvement of the quality of public services in all information agencies.

Mediation between Information Seekers and Librarians

Conversational analysis may be infrequently used as an explicit methodology in library and information science, but the ideas that underlie this family of discourse analysis are certainly present. A few examples amply demonstrate both the use and the utility of an understanding of the particular discursive practice of the reference interview. Catherine Sheldrick Ross (2003) summarizes some research findings. She relates one specific interaction:

> In the library visit study, a user who had asked for books about Richard Wagner returned to say that none of the books on Wagner contained the desired information. At that point, the librarian discovered belatedly that the user needed a plot synopsis for all of Wagner's operas and recommended an opera guide. The librarian admonished, "You could have saved a lot of time if you had just asked for that initially"—a good example of blaming the bad-guy user. (p. 40)

Ross provides a snapshot of a conversational analysis, a snapshot that probably would not have been possible had there not been a substantially accurate record of the transaction.

Sarah Anne Murphy (2005) offers a perceptive examination of reference exchanges as narrative texts. A combination of the patron's narrative (query), the librarian's efforts to clarify the patron's narrative (what she calls the "professional text"), and translation of the query into a systemic strategy (the "institutional text") embodies a hermeneutic event. Again, the examination relies on a record of high fidelity. This kind of discourse analysis frequently (and certainly in the case of the reference interview) has a practical focus, an aim of improving communicative effectiveness. Murphy says that "An awareness of the interactive texts exchanged during the reference narrative may also assist librarians in steering patrons away from a false-focus" (2005, p. 251). There can, then, be an educational benefit to such analysis. What Murphy finds carries implications for others studying reference interviews. Melissa Gross (1999) points out that, especially in educational settings, a patron's query may actually emanate from someone else. The imposed query may be a teacher's assignment, for example. An interview is needed for the imposed query to be correctly identified. While she does not advocate it directly, a discourse analysis of this exchange can help us understand questioners' articulations and ways librarians identify and respond to imposed queries. The interpretive examination of such queries can be very informative.

Many librarians are using technology to make reference services more accessible, and some of the services emulate real-time chats. Discourse analysis can be employed to examine these kinds of exchanges, but since they are not oral, there are differences that should be accounted for. For one thing, the computer-mediated communication that typifies online reference services may be a hybrid of spoken and written language. Discourse analysis of online reference services is simplified somewhat by the fact that exchanges tend to be dyadic. If one were to examine multi-user chats it would be difficult to follow some paths, since one person's response to a posting may be separated from the original by one or more other postings. In a traditional reference interview a librarian would have the benefit of nonverbal, as well as verbal, cues. Hesitation, apparent reluctance or confusion, and other phenomena might be communicative. In the absence of the nonverbal and the other oral aspects of messages, an analyst is left with a textual record of an exchange. Online reference services are becoming sufficiently common that careful examination of this form of communication is warranted for two basic reasons: (1) effectiveness of the service depends on a full understanding of the efficacy of the exchange (are the agents comprehensible to one another; is the medium adequate to the task of communicating questions and answers of all types; are responses accurate and complete), and (2) the nature of discursive patterns may present some particular challenges (the time required to type questions and responses may affect the cognitive-linguistic structures; the shortcuts that some people may take in their messages may necessitate longer series of questions and answers to insure clarity). Online exchanges share some characteristics

with telephone conversations but include differences in kind that render analysis unique. Jana Ronan (2003) provides a succinct illustration of some of the challenges that online reference presents from a discursive point of view. She recognizes the limitations to any conversational analysis of online transactions given the nonoral and nonaural restrictions. Her recommendations are primarily prescriptive, but an analysis could be employed to examine specific opportunities and inhibitors in an online exchange. She says, "Chat interviews often take longer, because questions that would be ambiguous at the reference desk may be even more confusing online, and there are no visual cues to add understanding" (p. 46). This phenomenon in particular is amenable to discourse analysis.

Discourse as Social Act

The second family of discourse analysis—the one that embraces the social, cultural, political, and other communicative acts as shown in Version 2 above—is also of importance to library and information science. Norman Fairclough offers a simple (possibly too simple) description of this family: "Critical approaches differ from non-critical approaches in not just describing discursive practices, but also showing how discourse is shaped by relations of power and ideologies, and the constructive effects discourse has upon social identities, social relations and systems of knowledge and belief, neither of which is normally apparent to discourse participants" (1992, p. 12). Michel Foucault is the theorist most frequently associated with this family of discourse analysis. Throughout the course of his life and work his ideas transformed a bit; I will address some aspects of both his archaeological and genealogical premises. These treat discourse first in the context in which it occurs and, second, with regard to more specific purposes. In his *Archaeology of Knowledge,* Foucault articulates a key question that situates inquiry: "The description of the events of discourse poses a quite different question: how is it that one particular statement appeared rather than another?" (1972, p. 27). The question highlights a concern of Foucault's that distinguishes his work from traditional intellectual history—that history tends to be sweeping and tends to embrace the totality of what is said on a topic or at a time. The archaeological process encompasses a focus on particulars. An archaeologist working on a dig examines not simply everything that can be found at a location but each artifact (including where it is found, how old it is, what is found near it, what might its uses have been and by whom, and other aspects of the artifact). Foucault expresses the difference between an archaeological approach and traditional intellectual history: "The analysis of the discursive field is oriented in a quite different way; we must grasp the statement in the exact specificity of its occurrence; determine its conditions of existence, fix at least its limits, establish its correlations with other statements that may be connected with it, and show what other forms of statements it excludes" (1972, p. 28).

Creating Relations and Ideologies

One of the things that distinguishes Foucault's approach is the promise that discourse not only reflects social relations and social action; it contributes to the construction of them. This second family usually examines formal discourse—texts, speeches, arguments, etc. Given these objects of study the second discourse analysis constitutes a study of ideologies (with "ideology" used not necessarily in any pejorative sense but as a formal articulation of a set of ideas or propositions and the rhetoric used to express them). This is one of the differences between the archaeological approach and intellectual history. The latter seeks to identify contradictions that can be resolved through unifying discourse. Archaeological analysis examines contradictions as they occur and as they are and not as problems to be solved or obstacles to be overcome (see Foucault, 1972, p. 151). Archaeological discourse analysis is not intrinsically concerned with what ought to be, in the sense of reaching the ultimate resolution to a puzzle; it is concerned with discursive practices as they are at a point in time. That point in time does have a past that has influenced the practice of the present. Also, that point in time is likely not to be unified; discursive practices may compete with one another, seek acceptance (some might say dominance), and embody the wills of the speakers. As Foucault points out in the "Discourse on Language," the competing practices, to be successful, rely on the nondiscursive actions typical of institutions:

> It is both reinforced and accompanied by whole strata of practices such as pedagogy—naturally—the book-system, publishing, libraries, such as the learned societies in the past, and laboratories today. But it is probably even more profoundly accompanied by the manner in which knowledge is employed in a society, the way in which it is exploited, divided and, in some ways, attributed (1972, p. 219).

Library and Information Science and Discursive Practice

At this point the work of Foucault probably sounds unendurably abstruse and abstract. He does, though, apply archaeological (and later his modified genealogical) analyses in specific environments—the prison, the hospital, science, and others. Moreover, Foucault's structures of analysis have been applied in library and information science. One specific application may clarify the use of Foucault's ideas in an analysis of our field's discourse. Bernd Frohmann (1992) employs discourse analysis to investigate writings advocating the cognitive viewpoint in library and information science. Frohmann draws explicitly from Foucault and incorporates an archaeological approach in his examination. His debt to Foucault is apparent in his article; he urges that "we shift our focus away from disputes over the truth or meaning of theoretical proposals, towards the *existence* of LIS theoretical discourses, treating as data for investigation and analysis the ways in which key theoretical ideas are talked about. Such a shift would involve pursuing implications of the fact that theory itself is a social practice" (p. 367).

Frohmann also uses Foucault's later genealogical approach as well. Archaeology and genealogy are certainly not mutually exclusive; they do, however, exhibit somewhat different focal points. The genealogical approach more explicitly examines the ways that discourse tracks not merely objective knowledge claims but the social relations based in power that define "objectivity" and attempt to legitimate knowledge claims. As is discourse itself, power is imbedded in and imbued with social relations that exercise a formative force. Power, its use, and those who exist within power relations are all evolving products of a historical complex of social interaction and definition. One of the institutions Foucault studies, the prison, did not spring *sui generis;* it has been based in theories of discipline, punishment, and (much later) rehabilitation. The individual—in this case the prisoner—is an object of study, an object observed while the observer is unseen. Foucault uses Jeremy Bentham's diagram of the panopticon (a design that enables guards to see in all directions without themselves being seen) as a model of disciplinary structure. The model for the ideal prison is, on its face, far removed from the ideal library, but the panopticon is not only a design for prisons:

> The Panopticon . . . must be understood as a generalizable model of functioning; a way of defining power relations in terms of the everyday life of men. . . . [T]he Panopticon must not be understood as a dream building: it is the diagram of a mechanism of power reduced to its ideal form; . . . it is in fact a figure of political technology that may and must be detached from any specific use. (Foucault, 1977, p. 205)

Within the context of the panopticon we can revisit the discourse surrounding library building design, perhaps the designs of Carnegie libraries in particular.

Foucault's work is not without its problems, and some of the difficulties are evident in uses of his work in library and information science. One of the most important challenges in his writings is the claim, more prevalent in his earlier books, that he is doing excavation rather than interpretation. That is, Foucault has said that his program involves detailing what is said and where it comes from (historically). The goal is not without interest, but interpretation inevitably enters into analysis. Hubert Dreyfus and Paul Rabinow (1983) describe the problem of both the archaeological and the genealogical approaches. They say, "This oscillation between description and prescription has revealed an even deeper instability concerning the status of serious meaning. . . . When viewed from this perspective, Foucault's methodological problems bear a suspicious similarity to the tensions he finds in the anthropological doubles" (pp. 90–91). The act of examining involves some interpretation. Foucault himself realizes the need for interpretation as he delves deeper into institutions that were the focal point of his early work. The second major challenge that Foucault presents is his assertion that knowledge, since it is inherently a function of power, does

not really have any objective existence. As he states, if historical consciousness "examines itself and if, more generally, it interrogates the various forms of scientific consciousness in its history, it finds that all these forms and transformations are aspects of the will to knowledge: instinct, passion, the inquisitor's devotion, cruel subtlety, and malice" (1977, p. 162). This reductive claim, if true, would require that only power be analyzed; nothing else has meaning.

Even with the shortcomings, some of what Foucault has articulated is very useful for analysis of discourses that are not conversational. Official documents, speeches, etc. are public and accessible and, by their nature, they speak to large audiences. In Version 2 above the purposes of persuasion or of presenting a notion that can be accepted are expressed in brief. A complicating factor, acknowledged by Barbara Johnstone (2002) and indicated earlier, is that discourse is both a product of social relations *and* produces social relations, is both a product of language *and* gives form to language (p. 9). This factor is at the heart of a problematic that Foucault described:

> "Words and things" [the original French title of The Order of Things is Les mots et les choses] is the entirely serious title of a problem; it is the ironic title of a work that modifies its own form, displaces its own data, and reveals, at the end of the day, a quite different task. A task that consists of not—of no longer—treating discourses as groups of signs (signifying elements referring to contents or representations) but as practices that systematically form the objects of which they speak. (1972, p. 49)

It should be clear from the foregoing background on this family of discourse that this conception of discourse analysis is not the same as content analysis. Differences should not be construed as superiority per se; each method has strengths and weaknesses and each can be used to address particular questions regarding particular works. Content analysis relies on categorizations—usually a combination of a priori and emergent categorization—as an analytical tool. Content analysis also tends to focus on texts (or sometimes images) as they are, without extensive historical situating. In many cases the intent behind the use of content analysis is to provide a current state, or snapshot, of a set of works (for example, violence in young adult books or favorable/negative editorial responses to political action). Both the archaeological and the genealogical approaches of Foucault point to a central difference between discourse analysis and content analysis: discourse analysis addresses more than an utterance. It is aimed at speech (*parole*), inasmuch as speech is historically situated, occurs at a point in time, and is engaged in by numerous individuals. Speech, therefore, embodies epistemological, rhetorical, communicative, obfuscatory, political, cultural, and other intentions. These intentions are essential to discourse analysis, and specific speech may simultaneously embody multiple intentions. This

speech, to borrow partially from Foucault, addresses matters of knowledge (in a generic sense). That is, the speech is aimed at what we know, what we think we know, what we can know, what institutions want us to know, etc. The connection to knowledge is of special importance to us in library and information science, since both professional practice and disciplinary inquiry are concerned with knowledge (how it is constructed, recorded, communicated, and preserved).

LIS, Power, and the Shaping of Discourse

The second family of discourse analysis has clear utility for us, and it has been employed by some researchers to address specific matters in our field. Some examples of application can help illustrate the strengths of discourse analysis. The examples also point to the most persistent and least overt challenge relating to discourse analysis—discourse analysis is, itself, discourse. It is also a discursive practice that can be subject to all of the analytical apparatus that it employs. A challenge for any analyst is to recall the imbeddedness of the speech employed with the speech that is studied. We can begin with a paper entitled "Public Space, Public Discourse, and Public Libraries" by Colleen Alstad and Ann Curry (2003, sec. 3, para. 1). The topic they address includes several intentions and is amenable to a discourse analytical approach. In their abstract they write: "The traditional mission of the public library—supporting the self-education of the citizenry in order that they may become fully participating members in a democratic society—has been devalued of late in favour of popularizing the library to attract more users." This statement is knowledge-based in that it articulates a specific position regarding what the public library mission should be and what it has become. This is historically situated speech that has cultural and political intentions. Their abstract continues: "By supporting public discourse, the public library can begin to reinvigorate both the quality of public discourse and its traditional commitment to democratic ideals." The statement is prescriptive, indicating that what is to follow in the paper will be a strategic discursive practice.

This is not to say that there is no analytical purpose to their paper, but it is not archaeology in the Foucauldian sense. The genealogical approach of Foucault, however, and the ideas of the will to knowledge and power are present in their analysis. What is said about technology, for instance, indicates that there are a couple of effects on public space: "The first is the 'virtualization' of the public sphere that is best exemplified by online discussion groups but also occurs on radio and television. The second is the manipulation of public discourse by mass media and its reconfiguration as an entertainment commodity." Alstad and Curry use themes from a conference to show that attention is on helping public libraries discover "what library customers want," understand "customers' interests," and develop a "strategy for marketing our products" and "our competitive edge." The

discourse, they aver, is imbued with a questioning of the public-ness of the public library: "By treating the library as if it were just another commercial enterprise, the popularization movement dismisses political, social, and moral values in favor of economics." The authors do not cite Foucault, but they do mention Jürgen Habermas, who has repeatedly argued for a normative, rather than an analytical, approach to discursive practice.

Another example of this kind of discourse analysis is an article by Siobhan Stevenson (2001). She draws most heavily from the theoretical and methodological work of Fairclough, which focuses on the social uses and social effects of discourse that have political and ideological elements. Stevenson says, "The three dimensions of this ideologically oriented model include text, discourse practice, and social practice" (2001, p. 53). She then offers a close analysis of some key documents emanating from the Canadian government that led to the establishment of some "Community Information Centres." Here analysis finds an underlying articulation of a societal shift through the government reports that is sufficiently critical and formative to suggest a change of direction for Canadian government policy. She reports that, "In such a world, there is no need for social action or social change. Social concerns are reconfigured as individual problems requiring individual solutions" (p. 70). Her work, as is the case with Alstad and Curry, cannot help but be a part of an "order of discourse" (see many of Foucault's works). The emphasis here must be on *an* order of discourse. Stevenson's analysis fits into what has become an institutionalized set of practices. Fairclough offers a particular point of view on orders of discourse: "the structuring of discourse practices in particular ways within orders of discourse can be seen, where it comes to be naturalized and win widespread acceptance, as itself a form of (specifically cultural) hegemony" (1992, p. 10).

Herein is a major challenge to the second family of discourse analysis but not an insurmountable one. Bernd Frohmann (2001) stresses Foucault's observation on the materiality of discourse (as recorded communication). His emphasis on this aspect of discourse is important; it reminds us of the existence of a statement, a claim, an utterance, an argument as it becomes material at a point in time. That material statement both cannot be removed from that time (inasmuch as it was articulated then) and exist at subsequent points in time. Foucault's archaeological approach eschews interpretation of statements in favor of the examination of the material circumstances of their existence (and Frohmann reiterates this position). The material nature of discourse is, of course, essential to analysis—statements say things in specific ways as part of a social structure and have historical and rhetorical functions. For example, a theoretical statement (that is, an articulation of a theory about a certain thing) is situated in the history of prior theoretical statements and embodies an effort to persuade that this statement is in some way superior to its predecessors. A community within a particular academic discipline may assess the theoretical statement

according to its explanatory and predictive merits; the discourse analysis examines it in the context in which it is produced (and can include the community's assessment). To be more specific, the scientific statements of Trofim Lysenko would not be analyzed according to their empirical efficacy but according to the social state (Stalinist Soviet Union) that enabled them to be produced and employed.

Frohmann's work illustrates the challenge. In examining the theoretical role of the cognitive viewpoint he asks, "If we take [Alvin] Schrader's notion of linguistic fashion to heart, are we then not challenged at least to investigate the possibility that fashions in LIS theory are perhaps as firmly grounded as the mutations of cultural taste?" (1992, p. 367). By way of a methodological answer he suggests that "we shift our focus away from disputes over the truth or meaning of theoretical proposals, towards the existence of LIS theoretical discourses, treating as data for investigation and analysis the ways in which key theoretical ideas are talked about. Such a shift would involve pursuing the fact that theory itself is a social practice" (1992, pp. 19–20). So far, there is an analytical problem; "fashions" (which are social, political, ideological, etc.) can be examined for what they are, and the historical situatedness of discourse can be studied. At the end of his paper Frohmann writes:

> The conclusion of the analysis presented here is that the "user-centric" promise of the cognitive viewpoint is compromised by the ways in which its discursive resources are mobilised to integrate users firmly within a market system of information consumption as much outside their control as any other highly monopolised system of consumer product production and exchange. (1992, p. 384)

His statement about the mobilization of resources to loci within a market economy stems from the discourse analysis itself. But how is the promise of the cognitive viewpoint compromised? What strictly material facets of the discourse render the conclusion plausible? Stated differently, how is his conclusion possible without interpretation? I am by no means denigrating Frohmann's work; I am merely pointing out the scope of the challenge that discourse analysis faces.

Interpretation

Now, how might we respond to this challenge? For one thing, we should follow Foucault in examining instances of discursive practice as events occurring at points in time. This applies to Foucault's own writings as well. While he did say that discourse should be studied as it is and without interpretation, he did, in fact, engage in interpretation. At the very least, discursive practice is connected to institutions and systems of knowledge, and those connections must be discerned and described. If there exists a will to knowledge, the will has some rationale, is instituted in some way, and is simultaneously reproduced and exercised. Following Frohmann,

if the cognitive viewpoint can be said (by anyone) to be *the* fundamental theory of information science, then it must have become institutionalized somehow. The uncovering of how such a thing occurred is interpretive, is achieved by examining who said what when, and determining how competing discursive practices were not successful in creating a sustainable will to knowledge. Within professions, discourse is (as Foucault repeatedly observes) controlled by an array of institutional procedures, many of which are sub rosa and not accessible for analysis. It is possible, as the authors mentioned here demonstrate, to examine public statements for the purpose of exposing discursive structures that tend to dominate communication in a field. What is *not* accessible, however, is what is not public. For example, we have no way of analyzing papers submitted to journals but not published. We do not know what peer reviewers had to say about those rejected papers. By default, analysis is limited to what, by institutional and procedural practices, become public utterances.

The connection between discourses and institutions has been addressed briefly in library and information science. Mark Day (2002) examines discursive "fashions" in library and information science management–related literature. He describes a land of iterative relationships that leads to what may be called a helical phenomenon. He writes, "Management discourse, in addition to defining the nature of its core concepts such as the consumer, employee, manager, and professional, also defines the basic nature of the corporate capitalist environment within which these social roles are enacted" (p. 235). The definition of concepts and environments turns on itself and contributes to a definition of discourse, and so on. Ron Day also provides a description of the complicated interrelations of discourse and institutions: "The alliance between professional discourses and often conservative and dominant ideological and cultural forces is not just a result of . . . accidental class alliances. . . . Critical studies of professions need to reach out to a broader social and cultural context in order to understand professions as products of social forces other than themselves" (2000, p. 471). Both of these observations echo a statement by Foucault, which, while possibly extreme, indicates a defining characteristic of our discursive lives: "The fundamental codes of a culture—those governing its language, its schemas of perception, its exchanges, its techniques, its values, the hierarchy of its practices—establish for every man, from the very first, the empirical orders with which he will be dealing and within which he will be at home" (1970, p. xx).

To repeat, Foucault's words are extreme to the point of threatening determinism. Softening his stance, we can more readily agree that part of our identities is socially influenced. That influence extends to our discursive practices in different social situations. One is likely to speak differently in, say, a committee meeting than in a casual conversation with a colleague. The influence further extends to the language that is likely to be deemed

appropriate to certain settings. The specific setting may reflect fairly clearly defined power relations. The classroom may be one such power-laden setting. The teacher may speak from the authority of the position (which includes deciding who among the students may speak) and the authority of knowledge (which usually means that the teacher is more learned than the students). An examination of the discourse that occurs in a classroom would have to acknowledge these relationships (see, for example, Bourdieu and Passeron, 1994). Of course, the preceding example does not describe all of the discursive practices that occur in educational settings. A graduate seminar is based on a different assumption of power relations that allows for greater freedom, openness, and candor. An analyst is required to recognize the different situations in which discursive practices are enacted. The admission of the influences of the social situatedness on identity and the ways identity is expressed is another way to respond to the challenge of interpretation. The will to knowledge is manifest in institutions, but it is not reified in the institutions; examination of educational settings demonstrates variability within institutions.

Summary

The discussion here focuses on two families of discourse analysis; the families are different in kind and in purpose. There are, as has been noted, similarities between the two families. In each the emphasis is on discourse *analysis*—the examination of discourse as it occurs. In each there is an attempt to study the effects of the discourse—what it means within the context in which it occurs. The effects of the discursive practices are also a matter of interest. With the first family a purpose is to gauge the efficacy of linguistic exchanges aimed at accomplishing particular objectives (such as locating relevant information in a reference transaction). A part of that purpose extends to assessing the understandability of exchanges (whether one person understands what the other is saying). That objective may be achieved by examining the discourse to see if the participants demonstrate understanding or by examining actions that can reflect understanding of what is said. With the second family a purpose is to study the circuitous routes taken by what is said. Everything said exists within the entire body of what has been said and responds to, refutes, borrows, opposes, adopts, manipulates, ignores, appropriates, and buries what has been said. For this family of analysis, Foucault provides a guide by not providing a guide. That is, he does not offer an explicit method of study; he does, however, present a way (his own way) of digging through what has been written and said, observing practices that exist in the company of other practices.

As is true of anything that could be called a methodology, discourse analysis offers a way of seeing things, of envisioning what is happening and what has happened. Each family of analysis proposes a set of eyes and ears so that we may see and hear some particular things that we are looking and

listening for. The examination of a reference exchange in a library is possible if the analyst comprehends the situation of the exchange. A question comes from somewhere; it has a genesis and an evolution that continues until it is spoken. What the analyst hears is that last state, the moment the question is asked of a librarian. This is what the librarian hears as well. The analyst can then examine whether the librarian takes the question as it occurs in that last state or attempts to extract its source and development. The inquiry's results can be descriptive, but they can also contribute to a normative practice. The examination of discursive practice is possible if the analyst comprehends the situatedness of the practice, the arrangement of the practice in time, place, etc. As Radford reminds us, "like any statement, whether it be a book on the library shelf or a single sentence within this article, historical documents do not speak for themselves. Their significance lies in their place within a greater discursive formation, that is, in the ways they are combined and arranged with other documents/statements" (2003, p. 14). Both families of discourse analysis offer possibilities for understanding; neither offers a mere mechanism, a simple blueprint to follow. As is true of any fruitful method of study, discourse analysis enlightens through creativity and is anything but a hammer in search of a nail.

References

Alstad, C., & Curry, A. (2003). Public space, public discourse, and public libraries. *LIBRES, 13*(1). Retrieved February 22, 2006, from http://libres.curtin.edu.au/libres13n1/pub_space.htm.

Austin, J. L. (1975). *How to do things with words* (2nd ed.). Cambridge, MA: Harvard University Press.

Brown, G., & Yule, G. (1983). *Discourse analysis.* Cambridge: Cambridge University Press.

Bourdieu, P., & Passeron, J-C. (1994). Introduction: Language and relationship to language in the teaching situation. In P. Bourdieu, J-C. Passeron, & M. de Saint Martin (Eds.), *Academic discourse* (R. Teese, Trans.) (pp. 1–34). Stanford, CA: Stanford University Press.

Day, M. T. (2002). Discourse fashions in library administration and information management: A critical history and bibliometric analysis. *Advances in Librarianship, 26,* 231–298.

Day, R. E. (2000). Tropes, history, and ethics in professional discourse and information science. *Journal of the Society for Information Science, 51*(5), 469–475.

De Saussure, F. (1959). *Course in general linguistics.* New York: Philosophical Library.

Dreyfus, H. L., & Rabinow, P. (1983). *Michel Foucault: Beyond structuralism and hermeneutics.* Chicago: University of Chicago Press.

Eliot, T. S. (1971). *The complete poems and plays, 1909–1950.* New York: Harcourt, Brace & World.

Fairclough, N. (1992). *Discourse and social change.* Cambridge: Polity.

Foucault, M. (1970). *The order of things: An archaeology of the human sciences.* New York: Pantheon.

Foucault, M. (1972). *The archaeology of knowledge and the discourse on language* (A. M. S. Smith, Trans.). New York: Pantheon.

Foucault, M. (1977). *Discipline and punish: The birth of the prison* (A. Sheridan, Trans.). New York: Pantheon.

Frohmann, B. (1992). The power of images: A discourse analysis of the cognitive viewpoint. *Journal of Documentation, 48,* 365–386.

Frohmann, B. (2001). Discourse and documentation: Some implications for pedagogy and research. *Journal of Education for Library & Information Science, 42,* 13–26.

Grice, P. (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.
Gross, M. (1999). Imposed versus self-generated questions: Implications for reference practice. *Reference & User Services Quarterly, 39,* 53–61.
Gumperz, J. J. (1982). *Discourse strategies.* Cambridge: Cambridge University Press.
Johnstone, B. (2002). *Discourse analysis.* Malden, MA: Blackwell.
Murphy, S. A. (2005). The reference narrative. *Reference & User Services Quarterly, 44*(3), 247–252.
Radford, G. P. (2003). Trapped in our own discursive formations: Toward an archaeology of library and information science. *Library Quarterly, 73*(1), 1–18.
Ronan, J. (2003). The reference interview online. *Reference & User Services Quarterly, 43*(1), 43–47.
Ross, C. S. (2003). The reference interview: Why it needs to be used in every (well, almost every) reference transaction. *Reference & User Services Quarterly, 43*(1), 38–43.
Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language.* Cambridge: Cambridge University Press.
Stevenson, S. (2001). The rise and decline of state-funded community information centres: A textually oriented discourse analysis. *Canadian Journal of Information & Library Science, 26*(2/3), 51–75.
Stubbs, M. (1983). *Discourse analysis: The sociolinguistic analysis of natural language.* Chicago: University of Chicago Press.
Vološinov, V. N. (1973). *Marxism and the philosophy of language* (L. Matejka & I. R. Titunik, Trans.). New York: Seminar Pres.
Wooffitt, R. (2001). Researching psychic practitioners: Conversation analysis. In M. Wetherell, S. Taylor, & S. J. Yates (Eds.), *Discourse as data: A guide for analysis* (pp. 49–92). London: Sage Publications.

Other Works Consulted

Budd, J. M., & Raber, D. (1996). Discourse analysis: Method and application in the study of information. Information Processing & Management, 32, 217–226.
Frohmann, B. (2004). Deflating information: From science studies to documentation. Toronto: University of Toronto Press.
Howarth, D. (2000). Discourse. Buckingham: Open University Press.
Ross, C. S., & Nilsen, K. (2000). Has the internet changed anything in reference? The library visit study, phase 2. *Reference & User Services Quarterly,* 40, 147–155.

John M. Budd is Professor and Associate Director in the School of Information Science and Learning Technologies at the University of Missouri–Columbia. Prior to his arrival at the University of Missouri he taught in the library and information science programs at Louisiana State University and the University of Arizona. He has published in several journals, including Library Trends, Library Quarterly, the Journal of Documentation, College & Research Libraries, Social Epistemology, and others. His most recent book is *The Changing Academic Library,* published in 2005 by the Association for College and Research Libraries. His previous book, *Knowledge and Knowing in Library and Information Science,* published by Scarecrow Press, won the 2002 Highsmith Award for library and information science literature.

Research in Constructivist Frameworks Using Ethnographic Techniques

Kirsty Williamson

Abstract

This article discusses the undertaking of research using a constructivist philosophical framework and ethnographic techniques, which can include individual interviews, focus groups, observation, and questionnaires. It begins with a broad overview of social sciences research philosophy, discussing both positivism and interpretivism, before moving on to focus on the constructivist paradigm, which comes under the interpretivist umbrella. The section on ethnography follows and includes sampling, data collection techniques, and data analysis. Examples from the author's work are used to illustrate both philosophy and method. They are from the author's studies of the information-seeking behavior of, firstly, women with breast cancer and, secondly, online investors.

Introduction

This article discusses not only a research method, ethnography, and the techniques that are commonly associated with it but also the philosophical framework in which the method can be situated. It begins with a broad overview of social sciences research philosophy and moves on to discuss the constructivist paradigm specifically before detailing ethnographic method and techniques and providing examples from the author's work.

Philosophical Traditions of Research in the Social Sciences

In the broader context of research theory in the social sciences, there are two major philosophical traditions—positivist and interpretivist (sometimes

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 83–101
© 2006 The Board of Trustees, University of Illinois

written as "interpretive").[1] In a nutshell, these two traditions are based on different assumptions about the nature of reality. Positivists consider that, as in the field of science, knowledge can only be based on what can be observed and experienced. Key positivist tenets are therefore "measurement" and "objectivity," resulting in a focus on quantitative data. The associated style of reasoning is "deductive," where the argument moves from general principles to particular instances. Positivist research usually begins with theories and models, defines variables for study, and predicts their relationships through framing hypotheses that are then tested. Generalizations are eventually made. Common research methods are "experimental design," with its emphasis on cause and effect, and "survey," which must be carried out according to scientific principles. "Validity" and "reliability" are key constructs for positivist researchers. (Powell, 1997, pp. 37–42 discusses these concepts).

On the other hand the interpretivist philosophy, where the constructivist paradigm fits, takes a different view of reality. Interpretivism is a broad term that encompasses a number of different paradigms, all concerned with the meanings and experiences of human beings. Since the central tenet of interpretivism is that people are constantly involved in interpreting their ever-changing world, researchers who are interpretivists believe that the social world is constructed by people and is therefore different from the world of nature (Williamson, 2002a). They favor "naturalistic inquiry" (where field work usually takes place in a natural setting), embrace an inductive style of reasoning, and emphasize qualitative data. It is the use of constructivist frameworks that is discussed in this article.

There are good reasons for using the terms "positivist" and "interpretivist" for describing researchers who subscribe to the two distinctly different ways of viewing the nature of reality. One arises because of the ways in which some key research theorists, such as Denzin and Lincoln (2003), discuss the field of "qualitative research." While they emphasize its interpretive nature and would include the interpretivist paradigms and methods, theirs is a broad, historical conceptualization and is not synonymous with interpretivist research. They say that "qualitative research is a field of inquiry in its own right," crosscutting "disciplines, fields, and subject matters" (p. 3). Another comment is that, historically, it is surrounded by "a complex, interconnected family of terms, concepts, and assumptions . . . [that] include the traditions associated with foundationalism, positivism, postfoundationalism, postpostivism, poststructuralism, and the many qualitative research perspectives, and/or methods, connected to cultural and interpretive studies" (p. 3).

This means that the term "qualitative research," on its own, does not provide an indication of the ontological view of the researcher. This is not to disregard the existence of those who postulate that there are some aspects of life, although not all, that are measurable, at least at a particular point in time, and who favor the use of mixed methods. In this case, too, the philosophical underpinnings of research should not be ignored. As Greene

and Caracelli say, "there is merit in different paradigmatic traditions in that each has something valuable to offer to our understanding of our complex social world. If such differences are not attended to in practice, then the full potential of mixed methods inquiry will remain unfulfilled" (2003, p. 107).

Constructivist Frameworks

"Constructivism," one of several interpretivist paradigms, is concerned with the ways in which people construct their worlds. Constructivist researchers investigate constructions or meanings about broad concepts such as cultural values, or more specific issues or ideas, such as the possible ingredients of the dynamic, creative public library of the future and how to create it. There are two major constructivist approaches—one focusing on individual, personal constructions and the other on shared meanings that could be said to reflect social constructions.

In the case of personal construct theory, a key proponent was Kelly (1955), who believed people make sense of their world on an individual basis, that is, personally construct reality. Some later cognitive researchers in the information-seeking field are theoretically closest to this form of constructivism. They moved beyond study of external, observable behavior to try to understand individuals from their own points of view. For example, Dervin and Nilan emphasized the importance of individuality, arguing that "the seeming complexity of individuality can be addressed . . . in a completely satisfactory manner which fulfils every reasonable demand of scientific investigation" (1986, p. 16). Dervin herself noted that the individually focused construction of her well-known Sense-Making methodology has been the most common one among information researchers who have adopted it (Olsson, 2003).

The other major constructivist approach comes from social constructionists who place emphasis on people developing meanings for their activities together (that is, socially constructing reality), as analyzed in the famous book *The Social Construction of Reality* (Berger & Luckman, 1967). In the information-seeking field, the social constructionist approach came to the fore in the late 1990s as discussions of the limitations of the cognitive and "information transfer" approaches to research, dominant for so long, began to appear in the literature. Tuominen and Savolainen (1997) and Talja (1997) pointed out the advantages of social constructionism. They all favored discourse analysis because the "processual negotiation of meanings" (Tuominen & Savolainen, 1997, p. 82) through which social reality is built occurs through discourse. Although not claiming the label, one of the early social constructionist researchers in the field was Elfreda Chatman, whose work focused on the information-seeking behavior of different communities and groups in specific social environments, such as older women living alone in a retirement village (Chatman, 1991, 1992) and prisoners

(Chatman, 1999). According to Olsson (2003), Chatman's approach was heavily influenced by Berger and Luckman. Olsson himself used a social constructionist framework to explore how information behavior researchers construct the meaning and significance of the work of the author Brenda Dervin. He said that, in her more recent writings, "Dervin has sought to challenge the construction of Sense-Making as a theory solely concerned with individual problem-solving," placing emphasis on the "social/collective aspects of Sense-Making" (p. 32).

There is no reason why researchers cannot draw on more than one body of research theory to underpin their own research. Indeed, Bates (2002) suggested that the three major metatheories as discussed by Tuominen, Talja, and Savolainen (as cited by Bates)—"the information transfer" model (which, according to Bates, they equate with a classically scientific approach), the constructivist model, and the constructionist model—should not struggle for dominance, with each being superseded in its turn. "The very fact that we have at some point in human history, explored and learned much that is meaningful from these various metatheoretical perspectives should suggest that there may be a *valuable continuing role for all of them*" (Bates, 2002, p. 13; emphasis in original).

Constructivist Grounded Theory

A constructivist approach to grounded theory has now been developed. Charmaz (2003) says that, unlike the original grounded theory, first developed by Glaser and Stauss (1967), constructivist grounded theory is not "objectivist." It "recognises that the viewer creates the data and ensuing analysis through interaction with the viewed" (Charmaz, 2003, p. 273), and therefore the data do not provide a window on an objective reality. Thus, there is recognition that researchers' backgrounds will influence their interpretations of the data. They cannot avoid being influenced by "disciplinary emphases" and "perceptual proclivities" (Charmaz, 2003, p. 259). This means that, although every effort is made to look at "how 'variables' are grounded—given meaning and played out in subjects' lives" (Dawson & Prus, 1995 & Prus, 1996, as cited by Charmaz 2003, p. 273), there is acceptance that researchers shape their data collection and redirect their analysis as new issues emerge (p. 271).

Ethnography

Williamson's (2002a) book on research methods in the field of information management and systems includes two specific chapters about ethnography—one from a theoretical perspective (Saule, 2002) and another about ethnographic techniques (Bow, 2002). In the second of these chapters, Bow talked of ethnography as being most closely linked with participant observation. She compared Saule's definition of ethnography in the earlier, theoretical chapter with Minichiello, Aroni, Timewell, and Alexander's

(1990) definition of participant observation, pointing out how similar they are—with emphases on studying people in their everyday contexts, or by participating in social interactions with them with the goal of understanding them. According to Bow, there is no single way of undertaking an ethnography or doing participant observation, "although many texts read as though there is only one set procedure" (2002, p. 267). She further noted:

> Participant observation is one of the most flexible techniques or set of techniques for doing research . . . [It] not only potentially combines a number of techniques, such as interviewing, focus groups, observation, and questionnaires, but also has the flexibility to emphasise some techniques over others, and to leave some techniques out altogether—depending on the requirements and constraints of the research itself, such as time, money and resources which are available. (Bow, 2002, p. 267)

It is important to emphasize that, as with all interpretivist research, ethnography is flexible in terms of research design with researchers seeking "to be totally open to the setting and subjects of their study" (Gorman & Clayton, 1997, p. 38). Although there is planning involved in that a literature search and review should be undertaken in order to understand the topic and research questions and a data collection plan should be developed, the research design tends to be nonlinear and iterative (meaning that the various elements in the research are interwoven, with the development of one influencing decisions about the others). For example, data analysis is undertaken throughout the project, not just in the concluding stage. There are now many "cutting edge" forms of ethnography (examples can be found in Denzin & Lincoln, 2005).

Sampling

Qualitative (interpretive) research depends on small samples that are purposively or purposefully selected. Patton observed that "the logic and power of purposeful sampling lies in selecting *information-rich cases* for study in depth. Information-rich cases are those from which one can learn a great deal about issues of central importance to the purpose of research; thus the term *purposeful* sampling" (1990, p. 169).

As this quotation implies, purposive samples are also often premised on the concept of "theoretical sampling" as discussed by Glaser and Strauss (1967). Theoretical sampling means selecting subjects who represent the important characteristics that researchers consider of interest to the study. With this approach there is no compunction to sample multiple cases that do not extend or modify emerging theory (Pidgeon & Henwood, 1996).

Data Collection Instruments

As mentioned above, there is a range of ethnographic data collection instruments from which to choose. The interview technique is a frequent choice, most commonly using open-ended or semi-structured questions.

There is a great deal of advice about interviewing in the social sciences literature and also in the Williamson (2002b) chapter in her research methods book. If a semi-structured interview schedule is used, it should be piloted so that you can be sure that you will be collecting the kinds of data you need. Nevertheless, once again there is flexibility to adjust questions to encompass new perspectives, especially in the early stages of data collection.

With regard to other techniques, if observation is chosen, this should be formalized through the development of a schedule, or set of questions, to guide the observation. Articles in this issue of *Library Trends* give guidance on "observation." The construction of a good questionnaire is a complex process, but there is much advice available in good-quality research methods texts, including in Williamson (2002b).

Data Analysis

Williamson and Bow (2002) provide considerable detail about how to code qualitative data. There are many sources that are helpful for learning about the process (for example, Miles & Huberman, 1994; Huberman & Miles, 2002; Silverman, 2001). Whether it is done manually or with a computer program, such as NVivo, the principles are the same, although there is no strict set of rules. The following are a few basic steps, which need to be supplemented with further reading:

1. Transcribe the data so that you have it in printed form.
2. Read through the data, making notes or memos about key points.
3. Categorize or label passages of data according to content so that identically labelled or categorized data can be retrieved as needed. Categories are made up of a short title, a definition if needed, and the data that relate to the category. Initially categories are usually broad and are subdivided to be more precise as the analysis progresses.
4. Conceptually organize the categories. This should start early in the process and continue throughout. It means thinking about the similarities, differences, and relationships among the categories, preferably representing this pictorially as recommended by Miles and Huberman (1994).
5. Develop themes in preparation for the writing up of the research findings.

Examples of Constructivist Research Using Ethnographic Techniques

This author draws on both personal and social construct theory for her research, believing that it is important to capture both shared and individual meanings—the consensus and the dissonance—about information seeking and use. She attempts to portray the multiple voices or perceptions about

the study's focal issues through the findings of her research reports, where quotes from participants are liberally included. Although she attempts to be fully open to the ideas and responses of her participants, she does not claim that her research findings are objective "truth" but rather a construction resulting from an interaction of the researcher and research participants in keeping with the Charmaz (2003) approach. In her studies she uses ethnographic techniques that are well suited to constructivist frameworks as they provide opportunities for researchers to try to elicit the perceptions, meanings, and experiences of participants and provide rich descriptions of them. As mentioned above, these techniques include interviews, either with individuals and/or in focus groups, questionnaires, observation, and examination of documents. In some studies (for example, McGregor & Williamson, 2005; Williamson, 1997), all or most of these techniques were used. With other studies, only one or two techniques have been used, in which case the study is discussed not as an ethnography but as a study using an ethnographic technique or techniques. As Saule said, "all of the frameworks within interpretivist ethnographies utilise triangulation" (2002, p. 184) since use of multiple techniques and theoretical constructs encourages validation of an ethnographic text. Where only one or two techniques are used, it is very important to use the literature to provide support for the findings.

This article now proceeds to a detailed description of two studies of information-seeking behavior, both using constructivist frameworks and ethnographic techniques. The first focused on information seeking for breast cancer using one ethnographic technique; the second focused on information seeking for online investment using two ethnographic techniques. Both of these studies are mentioned as examples of how Williamson (2005), in the recently published article in *Theories of Information Behavior: A Researcher's Guide,* expanded her original ecological model of human information behavior through research in constructivist frameworks.

Example 1: Information Seeking for Breast Cancer

At the 2002 Information Seeking in Context conference, Williamson and Manaszewicz (2002) presented a paper about the first stage of a project called Breast Cancer Knowledge Online (BCKOnline), where the major goal was to provide quality, "tailored" breast cancer resources to meet the differentiated information needs of the breast cancer community.[2]

The researchers considered this first stage as a study in its own right and referred to it as the "Breast Cancer Information Needs and Seeking (BCINS) Study." The paper provided a critical overview of the research about breast cancer information needs; discussed the need for user-centered, contextual studies of the information needs of women with breast cancer; discussed the potential of the Internet to assist in meeting breast cancer information needs; and outlined the project's philosophy and method and key findings from the study. Included here are the philosophy and method, from a later

stage of the project, and a sample of findings focusing on just one theme—preferences for information format, content, and presentation.

Philosophy and Method

For the BCINS study, the researchers adopted an interpretivist/constructivist approach in an attempt to understand breast cancer patients' perceptions, values, beliefs, and the "meanings" they construct around the issue of information needs, information seeking, and knowledge integration. Both personal constructs or individual meanings (Kelly 1955) and social constructs or shared meanings (Berger & Luckman, 1967) were of interest in the research. Williamson and Manaszewicz (2002) took the view that, when people share the experience of a certain disease such as breast cancer within a particular society, it is likely that some shared meanings will emerge and that the patterns can be used to improve services such as information provision. The researchers therefore set out to discover the meanings that were shared by participants as well as those that were not (consensus and dissonance).

The first phase of the project, the user needs analysis, involved fifty-nine women who currently had breast cancer or had had it in the past. The sample was a purposive one, selected to represent various age groups, disease stages, time since diagnosis, educational levels, marital status, urban and rural locations, and ethnic backgrounds. The researchers recruited participants through breast cancer nurses in both the public and private sectors and through facilitators of health care centers and breast cancer support groups. In addition, a separate focus group of eleven breast care nurses was convened, and seven family members of women with breast cancer were interviewed.

A combination of individual interviews and focus groups was used, with the intention of minimizing the weaknesses and maximizing the strengths of these two different styles of interviewing (Williamson, 2002b). The strength of individual interviews, mostly used in the earlier stages of interviewing, is that they enable interviewers to gain confidence with their subject matter before needing to manage and coordinate the range of views that usually emerge in a focus group. Individual interviewees are also unaffected by the views of others and so the "band-wagon" effect, which can occur in focus groups, is not a problem. On the other hand, the interaction in focus groups can be powerful in stimulating ideas and fruitful discussion.

The focus groups were mainly based on individuals who had a particular cultural or contextual factor in common, such as ethnicity, rural residence, age group, or attendance at a particular support group, as advocated in the literature (see, for example, Krueger, 1994). The eleven breast care nurses all took part in the same focus group. All individual interviews and focus groups were undertaken using a semi-structured interview schedule with a predetermined list of very broad questions. With the permission of the participants, all interviews were audio-taped.

Of particular interest in the interviews was the information participants found most useful in the past; their preferred information formats; their overall impression of the quality of the information they have used; the extent to which they used the Internet to locate breast cancer information; and the gaps they perceived in information provision, which were very important given that the end product of the research is a portal leading to information that is tailored to specific needs and backgrounds of women with breast cancer (for example, geographic location, age, ethnicity, literacy level, and time of diagnosis). An overarching aim was to identify the groups of people and types of information that should be specifically targeted in an online resource.

The audiotapes of the interviews were transcribed by an experienced transcriptionist. Although the analysis did not constitute a grounded theory, it was influenced by the "constructivist grounded theory" approach of Charmaz, which "recognises that the viewer creates the data and ensuing analysis through interaction with the viewed" (2003, p. 273). While an attempt was made to represent all views in the analysis and presentations of findings, the researchers were aware that the analysis was affected by the fact that a "template" needed to be constructed so that it could be used to develop a portal to information "tailored" to the differentiated information needs of the breast cancer community. The analysis was a continuous process with the initial categories, determined after the first few interviews, being continually reassessed and expanded as more data were collected. There were many themes developed, including the one focusing on "preferences for information format, content, and presentation," the findings for which are presented below. A matrix of demographic information was also developed.

Example of Findings: Preferences for Information Format, Content, and Presentation

With regard to all information sources, including the Internet, participants were asked about their preferences for breast cancer information format, content, and presentation. In many cases, women expressed strong preferences. Sometimes at the same time, as pointed out by Williamson and Manaszewicz (2002), they encapsulated dissatisfaction with the information currently available, both in terms of content and delivery mechanisms, confirming numerous studies focusing on information about breast cancer in the literature (for example, Fallowfield, 2001; Jenkins, Fallowfield, & Saul, 2001; Girgis, Boyes, Sanson-Fisher, & Burrows, 2000). As one participant said:

> She gave me a whole wad of information. I was furious. It was basically a repeat of each other. It was extremely patronizing . . . but it didn't actually talk about what it was doing biochemically. I wanted the hard data . . . I wanted diagrams. For the first time last week I actually saw

> what invasive (lobular) carcinoma looks like as opposed to a ductal or different type of cancer.

This participant was only one of the many who, supporting Bader and Strickman-Stein's (2003) finding, expressed a need for visual information. Another example comes from a younger participant whose answer, in response to a question about the improvements that can be made in information provision, was: "I think more visual, video stuff. Because if you see things as well as hear and read them you tend to recall stuff more. When I was on the chemotherapy my blood count dropped and I had to inject myself, but while I was at the hospital they showed me a video and . . . I found [it] helpful because it's visual."

Women also expressed needs for different types of information content. The woman quoted above, who was frustrated with the lack of detail and biochemical information contained in many of the resources she used, also said she would like different types of information content. Another participant who felt similarly said: "I don't want the throwaway type of article. I want the deep scientific type that I can take in." On the other hand, others felt they would like information to be "simplified," as expressed by this participant: "It needs to be simplified. It needs to be accessible because it's something that we need to know as much as doctors need to know. And if there is a way of translating it into layman's terms, I think we have a right to know."

As Williamson and Manaszewicz (2002) noted, these findings indicate the limitations of the approach of applying readability formulae (as occurs in the field of education) to patient information materials in order to assess their efficacy and relevance to the target audience. Several studies (Berland, et al., 2001; D'Alessandro, Kingsley, & Johnson-West, 2001; Beaver & Luker, 1997) assume that patient education materials should be aimed at the eighth grade level or below; however, according to D'Alessandro, Kingsley, and Johnson-West (2001), most patient education materials are still written at the tenth grade level or higher. As Williamson and Manaszewicz (2002) pointed out, in fact neither level is appropriate to all information seekers. For example, two of the participants quoted in this section would require a higher than tenth grade level for their information. They would not be alone given that, in May 2000, 30 percent of the Australian population had completed tertiary education (Australian Bureau of Statistics, 2001).

From this brief section of findings, you will notice the different voices of participants and their diverse views. These are not neatly categorized and packaged results as would emerge from the analysis of a self-administered questionnaire. Rather, they bring multiple layers and nuances reflecting the complexity of humans with their varied experiences and perceptions of issues that affect them. You will notice, too, the implications of the findings as drawn out by the researchers and the use of the literature to add confirmation or further debate to the discussion. As mentioned above, when the results are

not triangulated through the use of a number of research techniques, it is particularly important to use the literature in this role. These points will be reinforced in the next example.

Example 2: Information Seeking for Online Investment

Kingsford Smith and Williamson (2004) reported the results of a small pilot study of information seeking by Australian online investors.[3] It looked at the ways in which online investors seek financial information, as well as information about the online investing process itself. This pilot study underpinned an application for funding from the principal funding body of universities in Australia (the Australian Research Council). The researchers were successful with their application and a major study is now underway. In the interim, the pilot study is significant because little is known about how investors seek information without the advice of a professional advisor. As Kingsford Smith and Williamson (2004) pointed out, while once it was possible to infer that most investors would act on the advice of their advisors, how investors make investment decisions in nonadvisory, direct execution circumstances, which apply to online investing, is much more opaque. The findings of the research were positioned in relation to theory and empirical research from the generic field of community information-seeking behavior, including reference to the practices of those seeking information through the Internet, as frequently occurs with online investors. Since one of the researchers is a professor of law, and the end goal was to consider the implications of the findings for regulation of online investing, the findings were also set in the context of some important bodies of legal and economic thinking about information, price formation, and investment decision making in financial markets.

Philosophy and Method

The philosophy and method were very similar to those of the breast cancer project. One difference was that Kingford Smith and Williamson (2004) placed particular emphasis on social constructionist theory, which emphasizes the development of shared meanings through social processes involving people, language, and religion. Quoting Schwandt, who postulated that "we do not construct our interpretations in isolation but against a backdrop of shared understandings, practices, language, and so forth" (2000, p. 305), Kingsford Smith and Williamson pointed out the range of cultural influences, both macro and micro, on each individual and the common needs and understandings they are therefore likely to share. As in the breast cancer study, there was an interest in common, this time online investing, which meant that participants were again likely to have at least some shared perceptions of needs for information and elements of information-seeking behavior in common. The researchers therefore took the approach that the patterns that emerge from shared meanings can be

used to improve services such as information provision for online investing. Thus, the particular interest was in the shared meanings of participants, without ignoring those that were not shared—once again consensus and dissonance.

As a small pilot study, this project was well suited to the use of ethnographic techniques framed within the constructivist paradigm. One of the strengths of this approach is that it provides for exploring and generating ideas and, as a concomitant, the serendipitous findings it often elicits through its empirical research. In this case the framework allowed Kingsford Smith and Williamson (2004) to elicit rich-picture, in-depth perspectives from the small sample to which they were restricted for this particular study.

The sample was again a purposive one, selected to suit the needs of a small pilot study. It included representatives of two organizations that provide online investment services, E-Trade and COMMSEC, as well as the regulator, Australian Securities and Investments Commission, together with ten individual investors, selected to provide a mix of ages, genders, and socioeconomic and education levels along with some online investing experience. Kingsford Smith and Williamson's (2004) article, where further information about the sample can be found, focused on the findings from the individual investors.

With regard to data collection from the individual investors, two ethnographic techniques were used: individual interviews and a questionnaire to collect demographic data as well as additional information about investing and information-seeking behavior, collected in table format. In one part of the latter, participants listed all their Internet transactions and activities, including "information seeking," and rated the frequency thereof. The other part of the table asked investors what sources they used for financial information, again with frequency ratings. The sources of information and advice listed were information from broker's site/Internet discussion site; Internet execution with broker advice (Internet advisory); telephone execution with broker advice (telephone advisory); face-to-face advisory; newspapers/journals; printed literature from share brokers/financial experts; and information or advice from friends and acquaintances. The filling in of the table was assisted by the interviewers, who discussed each option with the interviewee, thus gleaning extra details and insights during the process. Then the semi-structured interview, lasting from one to one and a half hours, followed. It explored reasons for investing online or for continuing with traditional forms of investment, as well as in-depth discussion of information sources and their advantages and disadvantages.

The principles of the data analysis were the same as those for the BCINS project with, once again, constructivist grounded theory providing a major influence. In this case the full set of transcripts was read by both researchers, who compared their interpretations on a continuous basis. In relying on

interpretations from more than one researcher, Kingsford Smith and Williamson (2004) were seeking to acknowledge the role of constructivist researchers as the primary instruments in the research process (Marshall & Rossman, 1999) and to reflect on the effect of our own roles as they influenced the research process (Lincoln & Guba, 1985).

The quantitative data, which were collected through the questionnaire and table, were analyzed by the Statistical Package for the Social Sciences (SPSS), with the analysis involving only frequency counts, that is, the number of participants in each age group or the number who engaged in information seeking (together with the frequency of that activity).

Below are two related sections of findings—concerning personal sources of information and social intercommunication for online investing. As you will see, again there is an emphasis, as far as possible, on allowing participants to speak for themselves by using quotations to illustrate the views expressed.

Example of Findings: Personal Sources of Information for Online Investment

Participants were asked about their use of personal sources of information for their online investing activities. Early, foundational studies in the broad field of community information, where personal investing is encompassed, indicate that personal sources of information, such as family, friends, and acquaintances, are widely used for community information (see, for example, Warner, Murray, & Palmour,1973; Williamson, 1978; Chen & Hernon, 1982). In the legal field, Shiller and Pound's (1989) study of individual and institutional investors found that word of mouth contacts were important. The findings of our pilot study (Kingsford Smith & Williamson, 2004) provide further confirmation: all our participants talked with others who were also investors—ranging from family members to flat mates, friends, and work colleagues. A few investors described how they trusted and respected the investment prowess of members of their family. A couple of older investors particularly said they discussed investing with children, specifically the actual process of using the Internet for investing, where they relied on their children's greater familiarity with computers. On the other hand, one of the online investors said: "For a start I have no family members in Australia so most of my information comes from friends, acquaintances or old colleagues."

Several participants reported discussing investing with work colleagues whom they had discovered shared the interest and even using down time at work to carry out online transactions while chatting with colleagues. Nevertheless, some participants were quite clear that their trust was limited to one or two people whom they respected as knowledgeable: "It's probably my son would be the only one I trust because he has got pretty advanced skills as to how to access things. He is a pretty quick thinker. I don't think I would trust some of the other relatives who have not got quite the same

background." And "I speak to Paul about it because quite often we'll do it together. But he's got different ideas and he's more into the charts and I'm less into that. He's got different ideas on what some stocks will do and sometimes we talk about it but we don't really listen to each other that much. He's the only other one I'd speak to."

Finally, several investor interviewees acknowledged that information gained from friends and acquaintances was likely to be lacking in a significant respect. Investors were sensitive about talking about losses, and investor friends were also sensitive in probing them about losses. This means that information gained from these informal sources may be skewed in favor of good news and omit bad news. Asked what the reactions of friends were to the losses of the 2000 market crash, one investor echoed the comments of several others: "I think a lot of them just were silent. They really didn't say much and I didn't ask them. . . . If they had won a whole heap they would probably tell you."

The findings about the significance of personal sources are reinforced in the next section, which examines an interesting way in which personal sources were involved in individuals' investing activities.

Social Intercommunication and Investment Information Seeking

Despite the findings of the Shiller and Pound (1989) study, which had shown a considerable amount of interpersonal communication regarding information in the investing field, Kingsford Smith and Williamson (2004) were not prepared to find the level of social intercommunication encountered from the first interview. Despite the fact that very few reported using chat rooms or bulletin boards, there was a variety of social intercommunication. It ranged from casual conversations to regular semiformal meetings in pubs and coffee shops, to more formal discussion groups with a common interest in investing, and on to investor clubs in which members contributed to a common fund to learn from making actual investments. At the most structured, there are associations such as the Australian Shareholders Association, which conducts regular meetings with a formal agenda but has an opportunity for socializing afterwards, and the Securities Institute of Australia, which has a formal program of securities industry education and training. The Australian Stock Exchange also conducts seminars on issues of current interest to investors, which some of our investors reported attending.

One strong observation is the extent to which many investors see investing as a leisure activity. For example, one investor said: "The people that do it [online invest] every day, . . . a lot of them will basically communicate with their friends. They will be on the phone. It's like playing a video game where everyone's connected. So they ring each other 'What are you doing?'

They're networks. They go to seminars. They do all that sort of stuff." Another investor reported: "I see it as a bit of a hobby to fill in time."

This sense of investing as a hobby or leisure activity segued into the activity of attending informally organized social occasions in which investment trends are discussed and there is commonly a speaker who makes an informal presentation on a topic of current interest. Three of Kingsford Smith and Williamson's (2004) investors had recently, or still did, attend such occasions regularly, and several others reported having heard about them. As one said: "But there's quite a lot of seminars where you can go and the room is full. I've been to seminars and there's 40 or 50 people . . . the last time we met was at a pub . . . they have a few drinks . . . they have slides and they were talking about options trading . . . it's just a private group and they network with each other and if they meet someone new who's interested they say: 'Come along.'"

Another investor, who is female, reported that the shareholders' group she belonged to had been going about six years and was mainly male. She said: "People are very open and free about information I find. There is no covetousness very much. . . . One probably accumulates information by osmosis as much as anything. And so I think you always pick up some little thing."

Another investor who had actually convened one of these investors' social groups for a while described a family group that operated an investment club. She spoke of how the family "talk[s] about shares, and they put in so much and actually invest it. But that's purely for their family only."

Again, an investor reported a similar formation in an informal all male investor club: "We put up $500 each and invest it and play with it in the investment club. . . . Its attitude is to trade in all the shares you normally wouldn't trade in and its idea is to discuss things that people might know about . . . It's mainly there for that purpose, learning." The investor described this activity as fun, even though the investment club was making a loss and had recently called on all members for a top-up in funds!

Despite this widespread social activity focusing on investment, there was also a variety of views about how influential (in the investor's own perception) the social intercommunication was on the investor's own investment decisions. One reaction from the day trader to the idea of investors' groups was: "No, definitely not. . . . I want to make it purely the charts [that] generate what I buy, not people at all." The female member of the mostly male investment group warned us: "I find it really interesting to listen and to take some ideas, just to weigh them up, but I wouldn't take it straight from there, no. I have done it once and I had my fingers burnt."

From these sections of findings from the online investment pilot project, you will notice that, as with the BCINS study, the picture of personal source use for online investment is built up through the voices of the

study participants. This was followed (in the article) by considerable discussion by the researchers about the implications of these findings. In this example there are fewer references to the literature than appeared in the breast cancer project, despite the earlier advocacy that the literature should be used for the triangulation of findings. The reason is that, in contrast to the topic of breast cancer, online investing research in relation to information seeking is very much in its infancy. In the latter study, though, a second technique was used—the questionnaire and table of investing and information-seeking practice—thus providing another source of data to add depth to the findings. Nevertheless, given the exploratory nature of the project, the findings can only be tentative and require further investigation, which indeed is happening.

Conclusion

As with any research method, the style of research described in this article has strengths and weaknesses. It is not suited to the investigation of all research questions, including those that depend on eliciting statistical data from large samples. The method is especially suited to the exploration of the "why" research questions—those requiring in-depth exploration. One disadvantage is that samples need to be small as the major techniques are time consuming and costly to use. Small samples appear unreliable to some critics. Generalizations beyond the sample are inadvisable without strong evidence from other studies and some would see this as another disadvantage. Nevertheless, it could be argued that generalizations are often tricky to make, even with positivist approaches. For example, even if the sample for a survey is randomly selected, supposedly meaning that generalizations can be made to a population, the response rate may be low, thus calling into question the representativeness of the sample where the participation of respondents has depended on self-selection.

Another disadvantage that would be perceived by positivists is the apparent discursiveness of the answers from participants, which often do not fit neatly into easily managed categories. Interpretivists would counter this second point by pointing out that, in positivist studies such as surveys, people's views will often not fit neatly into the little boxes representing categories chosen by researchers.

The constructivist/ethnographic approach enables the meanings or perspectives of participants to be studied indepth and their particular words to be used to convey their meanings directly to the reader. Ways of thinking about issues, which may not have occurred to the researchers, are often revealed. Thus, the complexities of the real world have some chance of emerging.

Notes

1. Other terms to describe these philosophies are "paradigms" and "epistemologies." Williamson (2002a) discusses the fluidity of terminology used in research. Case (2002, pp. 131–155) provides a useful discussion, highlighting the diversity and problems of terminology, as well as other issues of conceptualization of the research landscape.
2. This research was funded by an Australian Research Council Linkage Grant (2002–3) and a contribution from BreastCare Victoria, an initiative of the Victorian Department of Health and Human Services. The Breast Cancer Action Group Inc. (Victoria) was also an industry partner. Chief Investigators of the BCKOnline project were Professor Sue McKemmish, Head of the School of Information Management and Systems (SIMS), Monash University; Associate Professor Frada Burstein, SIMS; Associate Professor Julie Fisher, SIMS; Dr. Kirsty Williamson, SIMS; Ms. June Anderson, SIMS; and Ms. Sue Lockwood, the Breast Cancer Action Group Victoria (BCAG). Other personnel include Research Fellows Rosetta Manaszewicz (SIMS & BCAG) and Fiona Ross (SIMS); research students Pooja Malhotra, Jane Moon, and Chan Cheah; and programmer Sergio Viademonte.
3. The pilot study was funded by small grants from the Faculties of Law and the School of Information Studies at Monash University, Victoria, Australia. The major project, titled "One Day, We'll All Invest This Way! Regulating Online Investment," is funded by a three-year Australian Research Council Discovery Grant, with the Chief Investigators being Professor Dimity Kingsford Smith, now in the Faculty of Law at the University of NSW, Dr. Kirsty Williamson of the Monash University Caulfield Campus of Information Technology and the School of Information Studies at Charles Sturt University, and Professor Stephen Bottomley of the Law Faculty at the Australian National University.

References

Australian Bureau of Statistics. (2001). *Census of population and housing*. Canberra: AGPS. Retrieved October 11, 2003, from http://www.abs.gov.au/Ausstats/abs%40census.nsf/4079a1bbd2a04b80ca256b9d00208f92/7dd97c937216e32fca256bbe008371f0!OpenDocument#LevelEducation.

Bader, J. L., & Strickman-Stein, M. (2003). Evaluation of new multimedia formats for cancer communications. *Journal of Medical Internet Research, 5*(3), e16.

Bates, M. J. (2002), Toward an integrated model of information seeking and searching. *New Review of Information Behaviour Research, 3,* 1–15.

Beaver, K., & Luker, K. (1997). Readability of patient information booklets for women with breast cancer. *Patient Education & Counselling, 31,* 95–102.

Berger, P. L., & Luckman, T. (1967). *The social construction of reality: A treatise in the sociology of knowledge*. New York: Anchor Press.

Berland, G. K., Elliott, M. N., Morales, L. S., Algazy, J. I., Kravitz, R. I., Broder, M. S., et al. (2001). Health information on the Internet: Accessibility, quality and readability in English and Spanish. *Journal of the American Medical Association, 285*(20), 2612–2621.

Bow, A. (2002). Ethnographic techniques. In K. Williamson (Ed.), *Research methods for students and professionals: Information management and systems* (2nd ed., pp. 265–279). Wagga Wagga, NSW: Centre for Information Studies, Charles Sturt University.

Case, D. O. (2002). *Looking for information: A survey of research on information seeking, needs and behavior*. London: Academic Press.

Charmaz, K. (2003). Grounded theory: Objectivist and constructivist methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of qualitative inquiry* (2nd ed., pp. 249–291). Thousand Oaks, CA: Sage Publications.

Chatman, E. A. (1991). Channels to a larger social world: Older women staying in touch with the great society. *Library & Information Science Research, 42*(6), 438–449.

Chatman, E. A. (1992). *The information world of retired women*. Westport, CT: Greenwood Press.

Chatman, E. A. (1999). A theory of life in the round. *Journal of the American Society for Information Science, 50*(3), 207–217.

Chen, C-C., & Hernon, P. (1982). *Information seeking*. New York: Neal-Schuman.

D'Alessandro, D. M., Kingsley, P., & Johnson-West, J. (2001). The readability of pediatric patient education materials on the World Wide Web. *Archives of Pediatric and Adolescent Medicine, 155,* 807–812.

Denzin, N. K., & Lincoln, Y. S. (2003). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of qualitative inquiry* (2nd ed., pp. 1–45). Thousand Oaks, CA: Sage Publications.

Denzin, N. K. & Lincoln, Y. S. (Eds.) (2005). *The Sage handbook of qualitative research* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Dervin, B., & Nilan, M. (1986). Information needs and uses. *Annual Review of Information Science and Technology (ARIST), 21,* 3–33.

Fallowfield, L. (2001). Participation of patients in decisions about treatment for cancer. *British Medical Journal, 323,* 1144.

Girgis, A., Boyes, A., Sanson-Fisher, R. W., & Burrows, S. (2000). Perceived needs of women diagnosed with breast cancer: Rural versus urban location. *Australian and New Zealand Journal of Public Health, 24*(2), 166–173.

Glaser, B. G., & Strauss, A. S. (1967). *The discovery of grounded theory.* Chicago: Aldine.

Greene, J. C. & Caracelli, V. J. (2003). Making paradigmatic sense of mixed methods practice. In A. Tashakkori & C. Teddue (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 91–110). Thousand Oaks, CA: Sage Publications.

Gorman, G. E., & Clayton, P. (1997). *Qualitative research for the information professional: A practical handbook.* London: Library Association.

Huberman, A. M., & Miles, M. B. (2002). *The qualitative researcher's companion.* Thousand Oaks, CA: Sage Publications.

Jenkins, V., Fallowfield, L., & Saul, J. (2001). Information needs of patients with cancer: Results from a large study in UK cancer centres. *British Journal of Cancer, 84*(1), 48–51.

Kelly, G. (1955). *The psychology of personal constructs.* New York: Norton.

Kingsford Smith, D., & Williamson, K. (2004). How do online investors seek information, and what does this mean for regulation? Theory, method and preliminary findings. *Journal of Information, Law & Technology,* 2. Retrieved February 22, 2006, from http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2004_2/kingsford-smithandwilliamson.

Krueger, R. A. (1994). *Focus groups: A practical guide for applied research.* Thousand Oaks, CA: Sage Publications.

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

Marshall, C., & Rossman, G. B. (1999). *Designing qualitative research* (3rd ed.). Thousand Oaks, CA: Sage Publications.

McGregor, J., & Williamson, K. (2005). Appropriate use of information at the secondary school level: Understanding and avoiding plagiarism. *Library & Information Science Research, 27*(4), 496–512.

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

Minichiello V., Aroni R., Timewell E., & Alexander, L. (1990). *In-depth interviewing: Researching research. A resource for social scientists and practitioner researchers.* Oxford: Blackwell.

Olsson, M. R. (2003). *The construction of the meaning and significance of an "author" among information behaviour researchers. A social constructivist approach.* Unpublished doctoral dissertation. University of Technology, Sydney.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage Publications.

Pidgeon, N., & Henwood, K. (1996) Grounded theory: Practical implementation. In J. T. Richardson (Ed.), *Handbook of qualitative research methods for psychology and the social sciences* (pp. 86–101). Leicester, England: British Psychological Society.

Powell, R. R. (1997). *Basic research methods for librarians* (3rd ed.). Norwood, NJ: Ablex.

Saule, S. (2002). Ethnography. In K. Williamson, *Research methods for students and professionals: Information management and systems* (2nd ed., pp. 177–193). Wagga Wagga, NSW: Centre for Information Studies, Charles Sturt University.

Schwandt T. (2000). Three epistemological stances for qualitative inquiry: Interpretivism, hermeneutics, and social constructionism. In N. K. Denzin and Y. S. Lincoln, (Eds.), *Handbook of qualitative research* (pp. 292–331). Thousand Oaks, CA: Sage Publications.

Shiller, R., & Pound, J. (1989). Survey evidence on diffusion of interest and information among investors. *Journal of Economic Behaviour and Organisation, 12,* 47–66.

Silverman, D. (2001). *Interpreting qualitative data: Methods for analysing talk, text and interaction.* London: Sage Publications.

Talja, S. (1997). Constituting "information" and "user" as research objects: A theory of knowledge formations as an alternative to the information-man theory. In P. Vakkari, R. Savolainen, & B. Dervin (Eds.), *Information seeking in context: Proceedings of an International Conference on Research in Information Needs, Seeking and Use in Different Contexts, 14–16 August, 1996, Tampere Finland* (pp. 67–80). London: Taylor Graham.

Tuominen, K., & Savolainen, R. (1997). A social constructionist approach to the study of information use as discursive action. In P. Vakkari, R. Savolainen, & B. Dervin (Eds.), *Information seeking in context: Proceedings of an International Conference on Research in Information Needs, Seeking and Use in Different Contexts, 14–16 August, 1996, Tampere, Finland* (pp. 81–96). London: Taylor Graham.

Warner, E. S., Murray, J., & Palmour, V. E. (1973). *Information needs of urban residents* (Report No. OEC-O-71-455, ED088464). Final report from the Regional Planning Council of Baltimore and Westat Inc. of Rockville, MD, to the U.S. Department of Health, Education and Welfare, Office of Education, Division of Library Programs.

Williamson, K. (1978). *Library Use and Information Needs in the City of Ringwood.* Melbourne, AU: Melbourne State College.

Williamson, K. (1997). Older people and barriers to public Internet access. *Internet Research: Electronic Networking Applications and Policy,* 7(3), 229–232.

Williamson, K. (2002a). *Research methods for students, academics and professionals: Information management and systems* (2nd ed.). Wagga Wagga, NSW: Centre for Information Studies, Charles Sturt University.

Williamson, K. (2002b). Research techniques: Questionnaires and interviews. In K. Williamson, *Research methods for students and professionals: Information management and systems* (2nd ed., pp. 235–249). Wagga Wagga, NSW: Centre for Information Studies, Charles Sturt University.

Williamson, K. (2005). Ecological theory for the study of human information behaviour. In K. E. Fisher, S. Erdelez, & E. F. McKechnie (Eds.), *Theories of information behaviour: A researcher's guide* (pp. 128–132). Medford, NJ: Information Today.

Williamson, K. & Bow, A. (2002). Analysis of quantitative and qualitative data. In K. Williamson, *Research methods for students and professionals: Information management and systems* (2nd ed., pp. 285–303). Wagga Wagga, NSW: Centre for Information Studies, Charles Sturt University.

Williamson, K. & Manaszewicz, R. (2002). Breast cancer information needs and seeking: Towards an intelligent, user sensitive portal to breast cancer knowledge. *New Review of Information Behaviour Research: Studies of Information Seeking in Context, 3,* 203–219.

Dr. Kirsty Williamson is director of a research group, Information and Telecommunications Needs Research, a joint initiative of Monash University and Melbourne and Charles Sturt University, Wagga Wagga, Australia. Since 1991 she has won a large amount of research funding enabling her to work as a researcher specializing principally in qualitative methods. Her principal research focus is information-seeking behavior, especially in relation to telecommunications. She has worked with a wide range of target groups but "older people" have been of particular interest and were the subject of her Ph.D. research.

Evaluation Research: An Overview

Ronald R. Powell

Abstract

Evaluation research can be defined as a type of study that uses standard social research methods for evaluative purposes, as a specific research methodology, and as an assessment process that employs special techniques unique to the evaluation of social programs. After the reasons for conducting evaluation research are discussed, the general principles and types are reviewed. Several evaluation methods are then presented, including input measurement, output/performance measurement, impact/outcomes assessment, service quality assessment, process evaluation, benchmarking, standards, quantitative methods, qualitative methods, cost analysis, organizational effectiveness, program evaluation methods, and LIS-centered methods. Other aspects of evaluation research considered are the steps of planning and conducting an evaluation study and the measurement process, including the gathering of statistics and the use of data collection techniques. The process of data analysis and the evaluation report are also given attention. It is concluded that evaluation research should be a rigorous, systematic process that involves collecting data about organizations, processes, programs, services, and/or resources. Evaluation research should enhance knowledge and decision making and lead to practical applications.

What Is Evaluation Research?

Evaluation research is not easily defined. There is not even unanimity regarding its name; it is referred to as *evaluation* research and *evaluative* research. Some individuals consider evaluation research to be a specific research method; others focus on special techniques unique, more often

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 102–120
© 2006 The Board of Trustees, University of Illinois

than not, to program evaluation; and yet others view it as a research activity that employs standard research methods for evaluative purposes. Consistent with the last perspective, Childers concludes, "The differences between evaluative research and other research center on the orientation of the research and not on the methods employed" (1989, p. 251). When evaluation research is treated as a research method, it is likely to be seen as a type of applied or action research, not as basic or theoretical research.

Weiss, in her standard textbook, defines evaluation as "the *systematic assessment* of the *operation* and/or the *outcomes* of a program or policy, compared to a set of *explicit* or *implicit standards,* as a means of contributing to the *improvement* of the program or policy" (1998, p. 4; emphasis in original). While certainly not incorrect, this definition, at least within a library and information (LIS) context, is too narrow or limited. Wallace and Van Fleet, for example, point out that "evaluation has to do with understanding library systems" (2001, p. 1). As will be noted later in this article, evaluative methods are used for everything from evaluating library collections to reference transactions.

Why Evaluate?

But before examining the specific techniques and methods used in LIS evaluation research, let us first briefly consider the question of why evaluation is important and then identify the desirable characteristics of evaluation, the steps involved in planning an evaluation study, and the general approaches to evaluation. With regard to the initial question, Wallace and Van Fleet (2001, pp. xx-xxi) and others have noted that there are a growing number of reasons why it is important for librarians and other information professionals to evaluate their organizations' operations, resources, and services. Among those reasons are the need for organizations to

1. account for how they use their limited resources
2. explain what they do
3. enhance their visibility
4. describe their impact
5. increase efficiency
6. avoid errors
7. support planning activities
8. express concern for their public
9. support decision making
10. strengthen their political position.

In addition to some of the reasons listed above, Weiss (1998, pp. 20–28) identifies several other purposes for evaluating programs and policies. They include the following:

1. Determining how clients are faring
2. Providing legitimacy for decisions
3. Fulfilling grant requirements
4. Making midcourse corrections in programs
5. Making decisions to continue or culminate programs
6. Testing new ideas
7. Choosing the best alternatives
8. Recording program history
9. Providing feedback to staff
10. Highlighting goals

"Over the past decade, both academics and practitioners in the field of library and information science (LIS) have increasingly recognized the significance of assessing library services" (Shi & Levy, 2005, p. 266). In August 2004 the National Commission on Libraries and Information Science announced three strategic goals to guide its work in the immediate future. Among those three goals was the appraising and assessing of library and information services.

Characteristics and Principles of Evaluation

Childers (1989, p. 250), in an article emphasizing the evaluation of programs, notes that evaluation research (1) is usually employed for decision making; (2) deals with research questions about a program; (3) takes place in the real world of the program; and (4) usually represents a compromise between pure and applied research. Wallace and Van Fleet (2001) comment that evaluation should be carefully planned, not occur by accident; have a purpose that is usually goal oriented; focus on determining the quality of a product or service; go beyond measurement; not be any larger than necessary; and reflect the situation in which it will occur. Similarly, evaluation should contribute to an organization's planning efforts; be built into existing programs; provide useful, systematically collected data; employ an outside evaluator/consultant when possible; involve the staff; not be any fancier than necessary; and target multiple audiences and purposes (Some Practical Lessons on Evaluation, 2000).

In an earlier work on the evaluation of special libraries, Griffiths and King (1991, p. 3) identify some principles for good evaluation that still bear repeating:

1. Evaluation must have a purpose; it must not be an end in itself
2. Without the potential for some action, there is no need to evaluate
3. Evaluation must be more than descriptive; it must take into account relationships among operational performance, users, and organizations
4. Evaluation should be a communication tool involving staff and users
5. Evaluation should not be sporadic but be ongoing and provide a means for continual monitoring, diagnosis, and change

6. Ongoing evaluation should provide a means for continual monitoring, diagnosis and change
7. Ongoing evaluation should be dynamic in nature, reflecting new knowledge and changes in the environment

As has been implied, but not explicitly stated above, evaluation often attempts to assess the *effectiveness* of a program or service. On a more specific level, evaluation can be used to support accreditation reviews, needs assessments, new projects, personnel reviews, conflict resolution, and professional compliance reports.

Types of Evaluation Research

Before selecting specific methods and data collection techniques to be used in an evaluation study, the evaluator, according to Wallace and Van Fleet (2001), should decide on the general approach to be taken. They categorize the general approaches as ad hoc/as needed/as required or evaluation conducted when a problem arises; externally centered, or evaluation necessitated by the need to respond to external forces such as state library and accrediting agencies; internally centered, or evaluation undertaken to resolve internal problems; and research centered, or evaluation that is conducted so that the results can be generalized to similar environments. Other broad categories of evaluation that can encompass a variety of methods include macroevaluation, microevaluation, subjective evaluation, objective evaluation, formative evaluation (evaluation of a program made while it is still in progress), and summative evaluation (performed at the end of a program). The *Encyclopedia of Evaluation* (Mathison, 2004) treats forty-two different evaluation approaches and models ranging from "appreciative inquiry" to "connoisseurship" to "transformative evaluation."

Evaluation Methods

Having decided on the general approach to be taken, the evaluator must next select a more specific approach or method to be used in the evaluation study. What follows are brief overviews of several commonly used evaluation methods or groups of methods.

Input Measurement

Input measures are measures of the resources that are allocated to or held by an organization and represent the longest-standing, most traditional approach to assessing the quality of organizations and their resources and services. Examples of input measures for libraries include the number of volumes held, money in the budget, and number of staff members. By themselves they are more measurement than true evaluation and are limited in their ability to assess quality.

Output/Performance Measurement

Output or performance measures serve to indicate what was accomplished as a result of some programmatic activity and thus warrant being considered as a type of evaluation research. Such measures focus on indicators of library output and effectiveness rather than merely on input; are closely related to the impact of the library on its community; and, as is true for virtually all evaluation methods, should be related to the organization's goals and objectives.

As was just indicated, one critical element of performance measurement is effectiveness; another is user satisfaction. In addition to user satisfaction, examples of performance/output measures include use of facilities and equipment, circulation of materials, document delivery time, reference service use, subject search success, and availability of materials. The Association of Research Libraries (2004) identified the following eight output measures for academic libraries: ease and breadth of access, user satisfaction, teaching and learning, impact on research, cost effectiveness of library operations and services, facilities and space, market penetration, and organizational capacity. One could argue that not all of those eight measures represent true performance or output measures, but they are definitely measures of effectiveness.

Impact/Outcomes Assessment

The input or resources of a library are relatively straightforward and easy to measure. True measurement of the performance of a library is more difficult to achieve, and it is even more challenging to measure impact/outcomes or how the use of library and information resources and services actually affects users. Rossi, Lipsey, and Freeman (2004) point out that outcomes must relate to the benefits of products and services, not simply their receipt (a performance measure). However, given the increasing call for accountability, it is becoming imperative for libraries to measure outcomes or impact. Indeed, "outcomes evaluation has become a central focus, if not the central focus, of accountability-driven evaluation" (Patton, 2002, p. 151).

Some authors use the terms *impact* and *outcome* synonymously; others see them as somewhat different concepts. Patton (2002, p. 162) suggests a logical continuum that includes inputs, activities and processes, outputs, immediate outcomes, and long-term impacts. Bertot and McClure, in a 2003 article in *Library Trends* (pp. 599–600), identified six types of outcomes:

1. Economic: outcomes that relate to the financial status of library users
2. Learning: outcomes reflecting the learning skills and acquisition of knowledge of users
3. Research: outcomes that include, for example, the impacts of library services and resources on the research process of faculty and students

4. Information Exchange: outcomes that include the ability of users to exchange information with organizations and other individuals
5. Cultural: the impact of library resources and services on the ability of library users to benefit from cultural activities
6. Community: outcomes that affect a local community and in turn affect the quality of life for members of the community

Matthews (2004, pp. 109–110), in his book on measuring public library effectiveness, identifies six categories of outcomes or benefits for public libraries. Those six categories, with examples, are as follows:

1. Cognitive results: refreshed memory, new knowledge, changed ideas
2. Affective results: sense of accomplishment, sense of confidence
3. Meeting expectations: getting what they needed, getting too much, seeking substitute sources
4. Accomplishments: able to make better-informed decisions, achieving a higher quality performance
5. Time aspects: saved time, wasted time, had to wait for service
6. Money aspects: the dollar value of results obtained, the amount of money saved, the cost of using the service

Impacts more relevant to academic libraries and their users include improved test scores, better papers, publications, increased class participation, etc. (Powell, 1995). A book by Hernon and Dugan (2002) considers outcomes for both academic and public libraries. The latter include getting ideas, making contact with others, resting or relaxing, and being entertained. Markless and Streatfield (2001) examine impact indicators for public, school, and academic libraries. Among their impact targets for school libraries are "improved quality and type of communication between learners and LRC staff" and "enhanced user confidence" (p. 175). Seadle (2003) notes that outcome-based evaluation is increasingly used for digital library projects.

Service Quality

Service quality, briefly defined, is "the difference between a library user's expectations and perceptions of service performance" (Nitecki, 1996, p. 182). As a concept, it dates back to at least the 1970s and has some roots in the total quality management (TQM) movement. TQM is characterized by the implementation of standards of quality, the encouragement of innovation, the measurement of results, and the taking of corrective actions as needed. TQM emphasizes the use of a team approach to maximizing customer satisfaction. A 1996 article by Pritchard provides an excellent overview of TQM, as well as other approaches to determining quality.

Quality is an elusive concept for which there is no commonly accepted definition, but the assessment of service quality did get a boost from earlier research from Parasuraman, Berry, and Zeithaml (see Nitecki, 1996). They

developed a conceptual framework, the Gaps Model of Service Quality, and a widely used instrument, SERV-QUAL, for measuring service quality. The Gaps Model incorporates the following gaps, as measured by the SERV-QUAL questionnaire:

1. The discrepancy between customers' expectations and managements' perceptions of these expectations
2. The discrepancy between managements' perceptions of customers' expectations and service-quality specifications
3. The discrepancy between service-quality specifications and actual service delivery
4. The discrepancy between actual service delivery and what is communicated to customers about it
5. The discrepancy between customers' expected services and perceived services delivered (Nitecki, 1996, p. 182)

The most visible current iteration of SERV-QUAL in the library field is known as LibQUAL+. LibQUAL+ was developed by faculty members of Texas A&M University in partnership with the Association of Research Libraries (ARL) and is part of ARL's New Measures Initiative. Over the past few years LibQUAL+ studies have been conducted by hundreds of libraries, including many large university libraries in the United States. These studies are intended for libraries "to solicit, track, understand, and act upon users' opinions of service quality" (*LibQUAL+*, 2003). Questions in the LibQUAL+ questionnaire address library staff, print and electronic resources, service hours, facilities, equipment, and document delivery and gather the data needed to calculate the gaps described above. However, according to Shi and Levy, "the current LibQUAL+ is not yet an adequately developed tool to measure and represent a dependable library services assessment result" (2005, p. 272).

Individuals wanting to know more about the use of service quality methods in academic libraries may wish to read a book by Hernon and Altman (1996). Other models of quality assessment from a British perspective are considered by Jones, Kinnell, and Usherwood (2000).

Process Evaluation

The second stage in Patton's (2002) continuum described in the section on impact/outcomes assessment was *processes* or activities. "A focus on process involves looking at *how* something happens rather than or in addition to examining outputs and outcomes" (p. 159). "Process data permit judgments about the extent to which the program or organization is operating the way it is supposed to be operating, revealing areas in which relationships can be improved as well as highlighting strengths of the program that should be preserved" (Patton, 2002, p. 160). Process evaluation focuses on "what the

program actually *does*" (Weiss, 1998, p. 9). It "is the most frequent form of program evaluation" (Rossi, Lipsey, & Freeman, 2004, p. 57).

Process indicators are somewhat similar to performance measures, but they focus more on the activities and procedures of the organization than on the products of those activities. For example, a process evaluation of an acquisitions department would be concerned with how materials are acquired and prepared for the shelf, not on how many books are ultimately used. In an academic library setting, process indicators might include staff training and development, delivery styles, knowledge of the curriculum, and participation in assignments and grading (Markless & Streatfield, 2001). In his book on public library effectiveness, Matthews (2004) places process measures in three categories: efficiency, staff productivity, and library information system activity. More generally speaking, a process evaluation "might examine how consistent the services actually delivered are with the goals of the program, whether services are delivered to appropriate recipients, how well service delivery is organized, the effectiveness of program management, the use of program resources, and other such matters" (Rossi, Lipsey, & Freeman, 2004, p. 57). And ultimately, the evaluator would want to know the extent to which programs and services were actually implemented. Patton (2002) even argues that "implementation evaluation" is a distinct method, and in many cases implementation information is of greater value than outcomes information (p. 161).

Benchmarking

One of the relatively recent approaches to measuring the performance of libraries and other organizations is benchmarking. Benchmarking tends to fall into the "total quality management" category. Benchmarking "represents a structured, proactive change effort designed to help achieve high performance through comparative assessment. It is a process that establishes an external standard to which internal operations can be compared" (Jurow, 1993, p. 120). The 2000 *Standards for College Libraries* describes benchmarking as the process of evaluating a library's points of comparison—inputs and outputs—against its peers and aspirational peers. There are several types of benchmarking, one of which is referred to as competitive or performance benchmarking. Performance benchmarking utilizes comparative data gathered from the same field or the same type of organization. The data are usually derived from analyses of organizational processes and procedures. Benchmarking can be used to establish best practices, identify changes to improve services, evaluate opinions and needs of users, identify trends, exchange ideas, and develop staff. Peischl (1995) points out that candidates for benchmarking include the services or products of an organization, internal work processes, internal support functions, and organizational performance and strategy.

Standards

According to Baker and Lancaster, "standards have an important role to play in the evaluation of library services . . . When applied to libraries, however, *standards* refers to a set of guidelines or recommended practices, developed by a group of experts, that serve as a model for good library service" (1991, p. 321). Some general types of standards, as identified by Baker and Lancaster (1991), include technical standards (for example, cataloging codes), performance standards, output measures, input measures, qualitative standards, and quantitative standards.

Quantitative Evaluation

Any evaluation method that involves the measurement of quantitative/numerical variables probably qualifies as a quantitative method, and many of the methods already examined fall into this broad category. Among the strengths of quantitative methods are the evaluator can reach conclusions with a known degree of confidence about the extent and distribution of that the phenomenon; they are amenable to an array of statistical techniques; and they are generally assumed to yield relatively objective data (Weiss, 1998, pp. 83–84).

Experimental methods usually, but not always, deal with quantitative data and are considered to be the best method for certain kinds of evaluation studies. Indeed, "the classic design for evaluations has been the experiment. It is the design of choice in many circumstances because it guards against the threats to validity" (Weiss, 1998, p. 215). The experiment is especially useful when it is desirable to rule out rival explanations for outcomes. In other words, if a true experimental design is used properly, the evaluator should be able to assume that any net effects of a program are due to the program and not to other external factors.

On the other hand, experimental methods are relatively weak in producing findings that can be generalized to other situations because they are usually conducted in rather controlled settings. Also, experiments tend to be used to test the effects of one component of a program at a time rather than the entire program. Another limitation of the true or randomized experiment is that it is not well suited for evaluating programs in their early stages of implementation. If the program changes significantly before outcomes are measured, it will be difficult to determine which version of the program produced what effects (Rossi, Lipsey, & Freeman, 2004).

Survey methods are often quantitative in nature but lack the experiment's ability to rigorously test the relationship between a program or service and its outputs or impact. Questionnaires and interviews, and observation to a lesser degree, represent the most commonly used survey data gathering techniques. Other quantitative methods covered by the *Encyclopedia of Evaluation* (Mathison, 2004) include concept mapping, correlation, cross-sectional design, matrix sampling, meta-analysis, panel studies, regression analysis, standardized tests, and time series analysis.

Qualitative Evaluation

As is true for basic research, qualitative methods are becoming increasingly popular. In fact, "the most striking development in evaluation in recent years is the coming of age of qualitative methods. Where once they were viewed as aberrant and probably the refuge of those who had never studied statistics, now they are recognized as valuable additions to the evaluation repertoire" (Weiss, 1998, p. 252). The *Encyclopedia of Evaluation* (Mathison, 2004) includes thirty-seven qualitative methods. They are appropriate, of course, when the phenomena being evaluated do not lend themselves to quantification. A qualitative method "tends to apply a more holistic and natural approach to the resolution of the problem than does quantitative research. It also tends to give more attention to the subjective aspects of human experience and behavior" (Powell & Connaway, 2004, p. 59). "Qualitative strategies can be particularly appropriate where the administration of standardized instruments, assigning people to comparison groups [in experiments], and/or the collection of quantitative data would affect program operations by being overly intrusive" (Patton, 2002, p. 191). In addition, they can provide

1. greater awareness of the perspective of program participants and often a greater responsiveness to their interests
2. capability for understanding dynamic developments in the program as it evolves
3. awareness of time and history
4. special sensitivity to the influence of context
5. ability to enter the program scene without preconceptions or prepared instruments, and to learn what is happening
6. alertness to unanticipated and unplanned events
7. general flexibility of perspective (Weiss, 1998, p. 253).

Qualitative methods do have their disadvantages as well, of course. Among them are the following:

1. Limited ability to yield objective data
2. Limited ability to produce generalizable results
3. Limited ability to provide precise descriptions of program outcomes
4. Not well suited for developing *specific* answers about the relationship of particular program strategies or events to outcomes (Weiss, 1998, pp. 85–86)
5. Often relatively labor intensive to conduct

Cost Analysis

Simple cost analysis is basically a descriptive breakdown of the costs incurred in operating an organization. Cost-related techniques more concerned with the assessment of whether monies are being spent in an optimal fashion usually fall into one of two groups—cost-effectiveness studies and

cost-benefit analysis. "The term 'cost-effectiveness' implies a relationship between the cost of providing some service and the level of effectiveness of that service . . . Cost-effectiveness analyses can be thought of as studies of the costs associated with alternative strategies for achieving a particular level of effectiveness" (Lancaster, 1993, p. 267). Some examples of cost-effectiveness measures include the cost per relevant informational resource retrieved, cost per use of a resource, cost per user, cost per capita, and cost by satisfaction level (Lancaster, 1993; Matthews, 2004).

Cost-effectiveness analysis can be seen as "a truncated form of cost-benefit analysis that stops short of putting an economic value on . . . outcomes [benefits] of programs" (Klarman, 1982, p. 586). "'Cost-benefit,' clearly, refers to a relationship between the cost of some activity and the benefits derived from it. In effect, a cost-benefit study is one that tries to justify the existence of the activity by demonstrating that the benefits outweigh the costs" (Lancaster, 1993, p. 294). A typical cost-benefit analysis involves determining who benefits from and pays for a service, identifying the costs for each group of beneficiaries, identifying the benefits for each group, and comparing costs and benefits for each group to determine if groups have net benefits or net costs and whether the total benefits exceed the total costs.

Types of cost-benefit analysis described by Lancaster (1993) are

1. net value approach: the maximum amount the user of an information service is willing to pay minus the actual cost
2. value of reducing uncertainty in decision making
3. cost of buying service elsewhere
4. librarian time replaces user time (that is, the librarian saves the user time by performing his or her task)
5. service improves organization's performance or saves it money.

Other kinds of cost analysis discussed by Weiss (1998) and Matthews (2004) include the following:

1. Cost-minimization analysis: seeks to determine the least expensive way to accomplish some outcome
2. Cost-utility analysis: considers the value or worth of a specific outcome for an individual or society
3. Willingness-to-pay approach: asks how much individuals are willing to pay to have something they currently do not have
4. Willingness-to-accept approach: asks individuals how much they would be willing to accept to give up something they already have
5. Cost of time

Organizational Effectiveness

The determination of the effectiveness of an organization has been identified as one of the objectives for some of the methods described above, and, indeed, it may be more properly thought of as an evaluation objective than an evaluation method. Regardless, it is a crucial element of organizational assessment and has received considerable attention in the professional literature. Rubin (cited by Wallace and Van Fleet, 2001, pp. 13–14) identifies a number of criteria for effectiveness at the organizational level and then describes several models for measuring organizational effectiveness. Those models and their "key questions" are as follows:

1. Goals: Have the established goals of the library been met?
2. Critical Constituencies: Have the needs of constituents been met?
3. Resources: Have necessary resources been acquired?
4. Human Resources: Is the library able to attract, select, and retain quality employees?
5. Open Systems: Is the library able to maintain the system, adapt to threats, and survive?
6. Decision Process: How are decisions made and evaluated?
7. Customer Service: How satisfied is the clientele with the library?

Program Evaluation Methods

In addition to the methods already identified, there are numerous other methods primarily used for social program evaluation. Readers interested in learning more about such methods are referred to the works on evaluation already cited above, including the article by Childers (1989), and to the table by King in Powell and Connaway (2004, pp. 57–58).

LIS-Centered Methods

Another approach to categorizing evaluation methods used in library and information science is according to the program, service, or resource to be evaluated. The book by Wallace and Van Fleet (2001), for example, has chapters devoted to the evaluation of reference and information services and to library collections (see Whitlatch, 2001 for an article on the evaluation of electronic reference services). Bawden (1990) presents a user-oriented approach for the evaluation of information systems and services. An earlier issue of *Library Trends* (Reed, 1974) has articles on the evaluation of administrative services, collections, processing services, adult reference service, public services for adults, public library services for children, and school library media services. Lancaster's 1993 text includes the evaluation of collections, collection use, in-house library use, periodicals, library space, catalog use, document delivery, reference services, and resource sharing. Most of these methods, however, actually employ techniques related to the more generic methods identified earlier in this article.

Planning the Evaluation Study

As has already been indicated, evaluation should be part of an organization's overall planning process and integral to the assessment of current services and resources, the development of strategies for change, and the monitoring of progress toward goals and objectives. Indeed, in order to be valid, an evaluation must reflect the organization's mission, goals, and objectives. In planning the evaluation of a specific program, the evaluator should first gather relevant background information. This activity might well include reviewing the professional literature, identifying professional standards and guidelines, and networking with colleagues. Next, the evaluator should decide what he or she actually wants to know, that is, focus the evaluation. This requires a determination of the purpose(s) of the evaluation specific to the program being examined. For example, the purpose may simply be to learn more about the program, or it may be to determine if the program is meeting its objectives.

After focusing the evaluation, decisions must be made about the overall design of the study, the method(s) to be used, and the measurements to be made. In other words, the evaluator must decide what must be measured, choose an evaluation method, select the data collection techniques to be employed, plan the construction and/or purchase of data collection instruments, plan the data analysis, develop a budget for the evaluation study, and recruit personnel. As is often the case in research studies, it is a good idea to utilize more than one method so as to increase the reliability and validity of the study and its findings. Haynes (2004, p. 19), for example, argues for mixed-method evaluation, which combines user-centered with system-centered paradigms and qualitative with quantitative methods. It is a good idea to write a thorough plan or proposal for the study at this time.

Weiss (1998) reminds us that the evaluator should also give careful thought to the best time to conduct the evaluation, the types of questions to ask, whether one or a series of studies will be necessary, and any ethical issues that might be generated by the study. Those and other planning points are succinctly represented in the following "evaluation action plan" suggested by Wallace and Van Fleet (2001, pp. 4–5):

1. What's the problem?
2. Why am I doing this?
3. What *exactly* do I want to know?
4. Does the answer already exist?
5. How do I find out?
6. Who's involved?
7. What's this going to cost?
8. What will I do with the data?
9. Where do I go from here?

Conducting the Evaluation Study

After planning the evaluation, it is time, of course, to conduct the study. That is, the evaluator is now ready to collect data or measure what needs to be measured; analyze the data; and report the findings. What follows is a brief overview of the steps in the evaluation process.

Measurement

"Measurement, in most general terms, can be regarded as the assignment of numbers to objects (or events or situations) in accord with some rule. The property of the objects which determines the assignment according to that rule is called *magnitude,* the measurable attribute; the number assigned to a particular object is called its *measure,* the amount or degree of its magnitude" (Kaplan, 1964, p. 177). More generally, measurement is any process for describing in quantitative values things, people, events, etc. Measurement by itself is not true evaluation, but it is one of the building blocks for quantitative evaluation. Common types of measures for library evaluation studies include number and types of users, number and duration of transactions, user and staff activities, user satisfaction levels, and costs of resources and services. They can be related to input, output, effectiveness, costs, etc.

It is critical that the measurement process and the measures be reasonably high in reliability and validity. Reliability refers to the degree to which measurements can be depended upon to secure consistent and accurate results in repeated applications. Validity is the degree to which any measure or data collection technique succeeds in doing what it purports to do; it refers to the meaning of an evaluative measure or procedure. The validity and/or reliability of measures can be affected by such factors as inconsistent data collection techniques, biases of the observer, the data collection setting, instrumentation, behavior of human subjects, and sampling. The use of multiple measures can help to increase the validity and reliability of the data. They are also worth using because no single technique is up to measuring a complex concept, multiple measures tend to complement one another, and separate measures can be combined to create one or more composite measures (Weiss, 1998).

Statistics

Many measures are in the form of statistics, which, in some cases, can be drawn from already existing sources of data. Types of statistics include administrative data, financial statistics, collections and other resources or inputs, use and other output/performance measures, outcomes, and staff and salary information. Sources of statistics include governmental agencies, professional associations, and other organizations such as state library agencies. Among the noteworthy sources of library-related statistics are the National Center for Education Statistics (NCES), American Library Association and its divisions (such as the Public Library Association's Public

Library Data Service and the Association of College and Research Libraries' Trends and Statistics series), Association of Research Libraries, and federal programs such as the Federal State Cooperative System and the Integrated Postsecondary Education Data System.

Data Collection Techniques

The evaluator must next select or design one or more data collection techniques that are compatible with the method(s) to be used and that are capable of gathering the necessary information. There are too many data collection techniques to consider here, but some of the relatively common techniques and instruments used for evaluation studies, as well as for other kinds of research, include the following:

1. Tests (standardized and locally developed)
2. Assessments by participants
3. Assessments by experts
4. Questionnaires (paper and electronic)
5. Interviews, including focus groups
6. Observation of behavior and activities
7. Evaluation of staff performance
8. Analysis of logs or diaries of participants
9. Analysis of historical and current records
10. Transactional log analysis
11. Content analysis
12. Bibliometrics, especially citation analysis
13. Use records
14. Anecdotal evidence

For information about many of these techniques, readers are referred to Powell and Connaway (2004) and Hernon and McClure (1990). For more information about techniques unique to evaluations of library and information use, readers may wish to consult earlier texts by Lancaster (1993) and Baker and Lancaster (1991). Westbrook's chapter in Powell and Connaway (2004), a chapter in Weiss (1998), and the book by Patton (2002) are among the sources of information about qualitative data collection techniques.

Analysis of Data

"The aim of analysis is to convert a mass of raw data into a coherent account. Whether the data are quantitative or qualitative, the task is to sort, arrange, and process them and make sense of their configuration. The intent is to produce a reading that accurately represents the raw data and blends them into a meaningful account of events" (Weiss, 1998, p. 271). The basic tasks of data analysis for an evaluative study are to answer the questions that must be answered in order to determine the success of the program or service, the quality of the resources, etc. Those questions

should, of course, be closely related to the nature of what is being evaluated and the goals and objectives of the program or service. In addition, the nature of the data analysis will be significantly affected by the methods and techniques used to conduct the evaluation. According to Weiss (1998), most data analyses, whether quantitative or qualitative in nature, will employ some of the following strategies: describing, counting, factoring (that is, dividing into constituent parts), clustering, comparing, finding commonalities, examining deviant cases, finding covariation, ruling out rival explanations, modeling, and telling the story. Evaluators conducting quantitative data analyses will need to be familiar with techniques for summarizing and describing the data (that is, descriptive statistics); and if they are engaged in testing relationships or hypotheses and/or generalizing findings to other situations, they will need to utilize inferential statistics.

Whatever the nature of the data analysis, however, it cannot substitute for sound development of the study and interpretation of the findings. Statistics can only facilitate the interpretation. In a quantitative study the analysis and interpretation usually follow the conduct of the study. In a qualitative study the data analysis is typically concurrent with the data gathering; "nor, in practice, are analysis and interpretation neatly separated" (Patton, 1987, p. 144).

The Evaluation Report

As part of the planning, the evaluator should have considered how and to whom the findings will be communicated and how the results will be applied. Weiss (1998, pp. 296–297) recommends that the typical report of a program evaluation include the following elements:

1. Summary of study results
2. Problem with which the program deals
3. Nature of the program: goals and objectives, activities, context, beneficiaries, staff
4. Nature of the evaluation
5. Comparison with evaluations of similar programs (optional)
6. Suggestions for further evaluation (optional)

A good report will be characterized by clarity, effective format and graphics, timeliness, candor about strengths and weaknesses of the study, and generalizability (Weiss, 1998), as well as by adequacy of sources and documentation, appropriateness of data analysis and interpretation, and basis for conclusions.

Conclusions

As was indicated above, evaluation research has been defined in a number of ways. It is viewed as a specific research methodology, as a type of study that uses standard social research methods for evaluative purposes,

and as an assessment process employing special techniques unique to the evaluation of programs. If treated as research, it is likely to be designed as applied or action research even though it may well use basic research methods. But generally speaking, all of the approaches to evaluation tend to share the following important commonalities: evaluation is a systematic process; it involves collecting data about organizations, processes, programs, services, and resources; it is a process for enhancing knowledge and decision making; and it is expected to lead to practical applications (Preskill & Russ-Eft, 2005, pp. 1–2). And finally, evaluation research should be conducted carefully and rigorously with consideration of many of the tenets that characterize good basic research.

References

Association of Research Libraries. (2004). *ARL new measures retreat; Transcript.* Washington, DC: Association of Research Libraries.

Baker, S. L., & Lancaster, F. W. (1991). *The measurement and evaluation of library services* (2nd ed.). Arlington, VA: Information Resources Press.

Bawden, D. (1990). *User-oriented evaluation of information systems and services.* Brookfield, VT: Gower.

Bertot, J. C., & McClure, C. R. (2003). Outcomes assessment in the networked environment: Research questions, issues, considerations, and moving forward. *Library Trends, 51*(4), 590–613.

Childers, T. (1989). Evaluative research in the library and information field. *Library Trends, 38*(2), 250–267.

Griffiths, J. M., & King, D. W. (1991). *A manual on the evaluation of information centers and services.* New York: American Institute of Aeronautics and Astronautics Technical Information Service.

Haynes, A. (2004). Bridging the gulf: Mixed methods and library service evaluation. *Australian Library Journal, 53*(3), 285–307.

Hernon, P., & Altman, E. (1996). *Service quality in academic libraries.* Norwood, NJ: Ablex.

Hernon, P., & Dugan, R. E. (2002). *An action plan for outcomes assessment in your library.* Chicago: American Library Association.

Hernon, P., & McClure, C. R. (1990). *Evaluation and library decision making.* Norwood, NJ: Ablex.

Jones, K., Kinnell, M., & Usherwood, B. (2000). The development of self-assessment tool-kits for the library and information sector. *Journal of Documentation, 56*(2), 119–135.

Jurow, S. R. (1993). Tools for measuring and improving performance. In S. Jurow & S. B. Barnard (Eds.), *Integrating Total Quality Management in a library setting* (pp. 113–126). New York: Haworth.

Kaplan, A. (1964). *The conduct of inquiry: Methodology for behavioral science.* San Francisco: Chandler.

Klarman, H. E. (1982). The road to cost-effectiveness analysis. *Milbank Memorial Fund Quarterly: Health and Society, 60*(4), 585–603.

Lancaster, F. W. (1993). *If you want to evaluate your library . . .* (2nd ed.). Champaign, IL: University of Illinois, Graduate School of Library and Information Science.

LibQUAL+ Spring 2003 survey results. (2003). Washington, DC: Association of Research Libraries.

Markless, S., & Streatfield, D. (2001). Developing performance and impact indicators and targets in public and education libraries. *International Journal of Information Management, 21*(2), 167–179.

Mathison, S. (Ed.). (2004). *Encyclopedia of evaluation.* Thousand Oaks, CA: Sage Publications.

Matthews, J. R. (2004). *Measuring for results: The dimensions of public library effectiveness.* Westport, CN: Libraries Unlimited.

Nitecki, D. A. (1996). Changing the concept and measure of service quality in academic libraries. *Journal of Academic Librarianship, 22*(3), 181–190.
Patton, M. Q. (1987). *How to use qualitative methods in evaluation.* Newbury Park, CA: Sage Publications.
Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.
Peischl, T. M. (1995). Benchmarking: A process for improvement. *Proceedings of the First International Conference on TQM and Academic Libraries* (pp. 119–122). Washington, DC: Association of Research Libraries
Powell, R. R. (1995). Impact assessment of university libraries. In *Encyclopedia of library and information science* (Vol. 55, pp. 151–164). New York: Marcel Dekker.
Powell, R. R., & Connaway, L. S. (2004). *Basic research methods for librarians* (4th ed.). Westport, CN: Libraries Unlimited.
Preskill, H., & Russ-Eft, D. (2005). *Building evaluation capacity: 72 activities for teaching and training.* Thousand Oaks, CA: Sage Publications.
Pritchard, S. M. (1996). Determining quality in academic libraries. *Library Trends, 44*(3), 572–594.
Reed, S. R. (Ed.). (1974). Evaluation of library service [Special Issues]. *Library Trends, 22*(3).
Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage Publications.
Seadle, M. (2003). Editorial: Outcome-based evaluation. *Library Hi Tech, 21*(1), 5–7.
Shi, X., & Levy, S. (2005). A theory-guided approach to library services assessment. *College & Research Libraries, 66*(3), 266–277.
Some practical lessons on evaluation. (2000). *Snapshots: Research Highlights from the Nonprofit Sector Research Fund,* No. 9.
Standards for College Libraries. (2000). Chicago: American Library Association.
Wallace, D. P., & Van Fleet, C. (2001). *Library evaluation: A casebook and can-do guide.* Englewood, CO: Libraries Unlimited.
Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
Whitlatch, J. B. (2001). Evaluating reference services in the electronic age. *Library Trends, 50*(2), 207–217.

Additional References

Abels, E. G., Kantor, P. B., & Saracevic, T. (1996). Studying the cost and value of library and information services. *Journal of the American Society for Information Science, 47*(3), 217–227.
Bertot, J. C., McClure, C. R., & Ryan, J. (2001). *Statistics and performance measures for public library networked services.* Chicago: American Library Association.
Childers, T. A., & Van House, N. A. (1993). *What's good? Describing your public library's effectiveness.* Chicago: American Library Association.
Everhart, N. (1998). *Evaluating the school library media center: Analysis techniques and research practices.* Englewood, CO: Libraries Unlimited.
Nelson, W. N., & Fernekes, R. W. (2002). *Standards and assessment for academic libraries: A workbook.* Chicago: Association of College and Research Libraries.
Nisonger, T. E. (2003). *Evaluation of library collections, access and electronic resources: A literature guide and annotated bibliography.* Westport, CN: Libraries Unlimited.
Smith, M. L. (1996). *Collecting and using public library statistics: A how-to-do-it manual for librarians.* New York: Neal-Schuman.
Van House, N. A., Lynch, M. J., McClure, C. R., Zweizig, D. L., & Rodger, E. J. (1987). *Output measures for public libraries* (2nd ed.). Chicago: American Library Association.
Van House, N., Weil, B, & McClure, C. (1990). *Measuring academic library performance: A practical approach.* Chicago: American Library Association.
Whitlatch, J. B. (2000). *Evaluating reference services: A practical guide.* Chicago: American Library Association.
Zweizig, D. (Ed.). (1994). *The tell it! manual: The complete program for evaluating library performance.* Chicago: American Library Association.

Ronald R. Powell is a professor in the Library and Information Science Program at Wayne State University, Detroit, Michigan. Prior to becoming a library and information science educator, he served as a university librarian and college library director. Dr. Powell has taught, conducted research, and published in the areas of research methods, collection development, bibliographic instruction, academic libraries, the measurement and evaluation of library resources and services, and education for librarianship. His publications include *Basic Research Methods for Librarians* (with Lynn Connaway), *Basic Reference Sources* (with Margaret Taylor), *Qualitative Research in Information Management* (with Jack Glazier), and *The Next Library Leadership* (with Peter Hernon and Arthur Young).

Encounters with the Library: Understanding Experience Using the Life History Method

Robert V. Labaree

Abstract

Drawing on the author's own research, this article explores the use of life histories as method and the ways in which this research can contribute to new understandings about the experiential relationships between libraries and users. The article is divided into four parts. Part one defines the essential elements of a life history research study. Part two describes how to design a life history research study. Part three examines ethical, methodological, and interpretive issues related to issues of organizational insiderness and internal validity and textual authority. The author concludes by outlining the potential benefits and pitfalls of using life histories and discusses how life history research, and qualitative research in general, can enrich and broaden our understanding of library science theory and practice.

Katheryn's profile is unusual for someone with aspirations of becoming a librarian. She recently graduated from the University of Southern California, earning a bachelor's degree in history with an emphasis in medieval society. The story of her pre-professional and pre-educational socialization into librarianship is somewhat unique because, as a woman in her early twenties, she made the deliberate decision to become a librarian many years before most of us would consider a career in librarianship. Katheryn explains, "Well, I've been going to libraries with my Mom since I was three, since she could bring me in there and be sure that I wasn't going to scream, and they were all good experiences."

Her mother was a volunteer in the local public library and Katheryn describes how being brought into "the back room of the library" where the photocopiers and other equipment resided made her curious about

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 121–139
© 2006 The Board of Trustees, University of Illinois

what people did there. However, as children, we are often exposed to the "back rooms" of other vocations. For example, a child may see cooks preparing meals at a restaurant, but that doesn't necessarily mean that he or she will develop an urge to be a chef. While I felt that it was important to understand Katheryn's socialization from a process and place perspective, I also wanted to understand what reflexive and affective factors may have contributed to her interest in becoming a librarian.

After listening to our first interview, I spoke with her a second time with the intent of getting her to think more about what drove her *desire* to become a librarian. In short, what triggered that moment from being in a library to wanting to become a librarian.

Author: In our first interview, you had stated that, in a broad sense, your desire at this point is primarily in public service. Has [being a student worker in the library] reinforced that desire?

Katheryn: In some ways, yeah. The excitedness, the weird, geeky excitedness of showing someone how to use [the resources]—that's *really* cool. I like that. And I like knowing where to point people to and having people who don't know where to find it, then having people go [there]. And I really like that. And I think that's the reason why I want to do public service, and, in fact, probably why I want to do more research and reference oriented librarianship than otherwise because there's that aspect of people actively looking for sources. And I can help them find them.

We both laughed out loud at that comment, but I knew exactly what she was talking about. I had felt the same thing as I began my own journey into this profession—this weird, geeky profession called librarianship.

Introduction

The above text is an excerpt from an ongoing, five-year life history research study that utilizes qualitative techniques of guided conversations, in-depth interviewing, and document analysis to understand the socialization experiences of seven young people who have made the deliberate decision to become librarians. Findings from this study will be used to develop new theories about occupational induction into librarianship that could inform better strategies of recruitment and retention. In capturing their life histories, the study attempts to not only document facts underlying how these individuals became socialized but to bring voice to the contextualized journeys of their own occupational induction experiences.

The decision to use life histories as a method to document processes of socialization into librarianship was based on three defining features of life history research described by Cole and Knowles (2001). First, life history research is intended to "advance understanding about the complex interactions between individuals' lives and the institutional and societal contexts in which they are lived" (p. 126). Similar to other service-oriented organizations governed by a commitment to educate their clientele, libraries possess a strong social connection to the people who utilize their resources

and services. It is within this domain that the decision to pursue a career in librarianship often takes root. A life history approach provides a method for documenting these experiences over time, placing them in proper social and cultural contexts, and executing a research project that helps answer questions about why someone might chose to become a librarian.

Cole and Knowles also describe life history studies as contributing "more just and dignified explorations and renderings of the human condition, that, in turn, lead to the enhancement of qualities and conditions under which lives are lived" (p. 126). A second defining feature of life history research is that it gives voice to the experienced life, particularly for those whose voices may be unheard or deliberately ignored or suppressed. Two subjects in my study are from underrepresented groups. Their stories of socialization are particularly important framed against current concepts of diversity and multiculturalism in librarianship (Honma, 2005) and as they relate to ongoing initiatives to recruit and retain minority librarians (Darby, 2005; Harralson, 2001).

Life history research also tells people's stories in their own words and, in this way, conveys a representation of human experience that draws readers into the interpretative process. Readers are invited to make meaning and to form judgments based on an interpretation of the text as it is viewed through their own realities. This is the third intention of life history research, according to Cole and Knowles (2001). By documenting the stories of seven individuals as they progress toward professional inclusion, I am attempting to construct stories of socialization that are relevant and accessible to the reader.

Contemporary research about the possible connections between the informal socialization of individuals prior to considering a career in librarianship and the eventual decision to become a librarian is empirically underdeveloped and largely anecdotal. While there are many methods a researcher could use to investigate this issue, I chose a life history approach because it provides an effective means of documenting, in depth and over time, individual stories of professional induction. As with other qualitative methodologies, researchers using a life history method must develop their studies based on good design, reflexive modes of implementation and analysis, and sound ethical principles.

The next section of the article will describe the essential design elements of a life history study. This is followed by an exploration of two critical methodological and ethical issues that may arise while conducting a life history study: negotiating organizational insiderness and the challenges associated with concepts of validity and textual authority. The article concludes by outlining the potential benefits and pitfalls of life history research and placing qualitative life histories within the larger milieu of library science research and practice.

Designing a Life History Study

A good life history study disrupts traditional assumptions about what is known or considered to be "the truth" and challenges the self-evident meaning of dominant culture language. This construct forces the reader to confront subjective perceptions of others (Goodson, 2001). However, the concept of the "life" in a life history study is somewhat misleading because an individual's entire biography is rarely the object of analysis (Kouritzin, 2000). Most life histories contextualize specific events or issues around the experienced lives of others. For example, Richie (2001) used life histories to investigate the challenges formerly incarcerated women faced when they returned to their communities. Grossman (1990) contributed to the literature on reforming teacher education by investigating the pedagogical content knowledge of English among beginning teachers and their emerging beliefs about classroom instruction. Sawyer (2005) used life histories to understand how various social institutions influence opportunities for active engagement in civic leadership by young people. In these and most other cases, life histories are purposefully bounded by the research question underpinning the study and do not attempt to document the entire life of an individual.

The concept of "history" in life history research refers to the practical aspects of how investigators must document the ways in which people experience the world. Unless an investigator can shadow the respondent wherever he or she goes, and can do so without influencing the collection of data and the interpretation of findings, the narrative stories in a life history are always a reflection of lives lived. As Jarvinen points out, "From the point of view of the present, there is no objective past in the history of individuals, institutions or societies. There is no past to be captured, understood and described in its pure essence. There is only a past—or plurality of pasts—constructed from the point of view of an ever-changing present" (2004, p. 47). From the standpoint of analyzing the data from life histories, the researcher must always understand that "With every new present, there comes a new past" (p. 47). Life histories always document the past and, therefore, findings represent perceptions of events as interpreted by the respondent at any given moment in the present.

Sampling and Identifying Data Sources

Consider the following: at a medium-sized, urban university, research conducted by the Office of Student Affairs demonstrates that significantly more first-generation students have difficulty adjusting to the academic rigors of college than those students with parents who attended college. In the library, first-generation students have been observed studying, but anecdotal evidence indicates that they rarely seek help at the reference desk or take advantage of the many services offered by the library. The librarians determine that one possible strategy to reduce stress and increase a sense

of self-efficacy among first-generation students is to develop programs to improve their knowledge of library resources and services. Unfortunately, no data exists to help us understand the information-seeking behavior of first-generation students and their utilization of library resources and services.

Designing a life history study around this research problem requires identification and selection of a representative sample of respondents and determining what types of additional sources could be used to triangulate the findings (Creswell, 1998). The underlying purpose is not to extrapolate broad generalizations or to formulate empiricist explanations of phenomena but to challenge existing assumptions, develop intimate familiarity with a specific issue, and, in this particular example, to gain insight into the experiences of first-generation college students as it is viewed from their own realities (Plummer, 2001).

Sample sizes in life history research are usually very small because gathering, recording, and interpreting the data can be intense and time-consuming. In addition, life histories rarely rely on methods of random sampling (Goodson & Sikes, 2001). According to Patton (1987; see also Morse, 1994), samples can be developed using one of four general methods:

1. *Extreme or intense case sampling.* Respondents are selected because their experiential characteristics "maximize the factors of interest" in a study (Morse, 1994, p. 229). Data from this approach is intended to clarify important factors relevant to the study.
2. *Maximum variety sampling.* This approach uses a heterogeneous group of respondents and documents commonalities among them. Data from this sampling method highlights cases of uniqueness or reveals shared patterns across the sample group.
3. *Critical case sampling.* This approach is used to ensure detailed, in-depth information on critical experiences. The purpose is to gather data on critical incidents that may inform other situations or events.
4. *Intensity sampling.* This approach emphasizes the selection of respondents because they are intimate authorities about a particular experience. Respondents are chosen because they possess a deep understanding of a particular issue or phenomena.

Once a sampling method has been determined, the identification of participants in a life history study generally involves a process of purposive discovery governed by convenience (the researcher has easy access to the respondent), opportunity (a chance meeting with someone willing to participate), or snowballing (a respondent identifies others who might be able to participate) (Goodson & Sikes, 2001).

Although conducting extensive interviews with respondents is the most common technique used to gather life history data, other biographical documents may be utilized. These can include autobiographies, memoirs,

diaries, personal journals, oral histories, electronic correspondence, and personal documents. This material can help to establish validity, understand what may have been omitted from memory, and verify factual information (Kouritzin, 2000; Roberts, 2002). When designing a life history study, it is important to consider any documents that illuminate and expand upon an individual's contextualized experiences. For example, in determining the information-seeking habits of first-generation students, permission to examine completed course assignments, library records, and email correspondence with professors could prove helpful.

Negotiating Participation and Access

The next step after identifying a sample of respondents is to negotiate access and participation (Goodson & Sikes, 2001). Because life history interviews are personal encounters that probe in depth the thoughts, feelings, and actions of others, a useful strategy for encouraging participation is to approach interaction with respondents from a social constructionist perspective (Shotter, 1993). Presented this way, new meaning and knowledge emerges in the form of a co-constructed journey of exploration. This dynamic transcends the basic dichotomy of the researcher and the researched to a more complex and sophisticated framework that acknowledges the context-dependent and communicative-driven interplay between the researcher, the participant, and the social worlds that they occupy.

Life history interviews may also elicit highly personal information or reveal illegal or deviant behavior (for example, "I only come to the library to download movies"). This raises important ethical issues. As a consequence, researchers have a responsibility to protect the privacy of anyone involved in the research project and to inform respondents of their rights as subjects of a research study (Johnson, 2002). These rights include being told the purpose and intended outcomes of the research study; knowing how and to what extent personal information will be protected; being told that they can ask questions or express concerns at any time before, during, or after the study; being told that they can withdraw from the study at any time; having a copy of any consent form used for the study; and knowing how, to whom, and in what form findings will be reported. For academic librarians conducting practitioner research, the rights of participants are governed by the institutional review boards of their school, and they must be followed very carefully (Pritchard, 2002).

Interviewing Techniques and Tactics

As noted, the most common technique for gathering life history data is to interview respondents. The goal is to create an in-depth profile of the respondent's life experiences relative to the research problem being investigated. From an organizational perspective, qualitative interviewing can be effective in evaluating library programs and services that focus on individualized outcomes; documenting and describing programmatic pro-

cesses; analyzing experiential variables among participants in a program; assessing trends in services or programs that are considered to be changing or evolving; understanding the underlying meaning of a service or program for participants; and identifying variations in design and implementation of programs at various sites (Patton, 1987, pp. 40–42). There are several comprehensive guides to doing qualitative interviews (Kvale, 1996; Rubin & Rubin, 1995; Seidman, 1998). Kvale, in particular, should be consulted for a detailed description of how to design a research study that relies on qualitative interviewing. However, the distinctive features of a life history project place greater emphasis on specific aspects of planning and implementing interviews with respondents. In general, these encompass four overlapping activities: being well prepared throughout the research process; utilizing unstructured, open-ended interview protocols; practicing active listening techniques; and conveying an understanding of the respondent's experience (Plummer, 2001).

Qualitative research interviews represent an active process of ongoing intellectual discovery. This means that a constant flow of new knowledge and meaning emerges from the examination of variables and their interrelationships identified from the data. Interviews require careful preparation and planning by the researcher. This is especially important with regard to developing an effective protocol that captures data relevant to the study's purpose while, at the same time, recognizing the need to schedule possible follow-up questions and to analyze relevant secondary documents that may help record and preserve context. The emergent nature of life history interviews also means that some questions must be adapted to each respondent's lived experiences since no two people experience events or interpret meaning in exactly the same way. Variables in experiences and interpretive meanings are important in developing a complete understanding of the phenomena under investigation. Finally, careful preparation also includes practical issues, such as purchasing a reliable tape recorder and scheduling a quiet place to conduct the interview. These issues may seem mundane, but they are vital in ensuring that each interview session maximizes the opportunity to reveal new data.

Careful planning is also important because life history interviews place a greater emphasis on unstructured and open-ended forms of inquiry. Generally defined as guided conversations (Cole & Knowles, 2001; Rubin & Rubin, 1995), the intention of a life history interview is to encourage a relationship with the respondent that is not "so blatantly purposeful that mutuality and authentic engagement is lost" (Cole & Knowles, 2001, p. 72). The conversation is guided because the purpose of the research is to conduct an intensive exploration of specific lived experiences; the purpose is not to develop a comprehensive biographical profile of the individual. Within this framework, the challenge for the researcher is to develop a guided conversation protocol that encourages opportunities to understand

as much as possible about relevant moments in a person's life but that is not so open-ended and unfocused that the questions generate an excessive amount of needless information.

The emergent nature of an unstructured, open-ended guided conversation requires active listening techniques in order to hear the underlying meaning of what is being said (Rubin & Rubin, 1995) and to obtain a deeper understanding of the knowledge possessed by others (Fiumara, 1990). During a life history interview, the researcher must be cognizant of the fact that new information may emerge at any time. However, the tape recorder is not a substitute for practicing good listening techniques. For example, in my life history study of newcomer socialization into the profession, "Roscoe" noted that one of his motivations for wanting to pursue a career in academic librarianship was witnessing what he described as "so many bizarre things around here and things that—that I don't think are right." Although this statement was made during an initial discussion about his growing professional interests, asking a series of follow-up questions was the key to unlocking the fact that being a library director was one of the initial factors influencing Roscoe's decision to pursue a degree in library science. It is important to note that, because life history interviews are the result of a discursively co-constructed journey of discovery, new meaning and new knowledge are not only revealed through the intersubjective relationship between the researcher and the respondent; the interaction itself becomes a contributor to revealing additional insight and understanding (Koschmann et al., 2005).

The fourth distinctive feature of a life history study that governs how interviews are designed and carried out is to show empathy to those we interview. Empathizing involves more than nonverbal cues of affirmation and acknowledgment made (often subconsciously) during an interview. To better understand the role of empathy in qualitative interviewing, I borrow from the conceptual work of Bondi (2003) and her exploration of power and positionality in feminist geography fieldwork. She notes that empathy is important in qualitative interviewing because it "enables the creation of interpersonal and intrapsychic spaces in which similarities and differences can be mobilised, expressed and explored" (p. 67). This is why practicing good listening skills is so important. An interview is a division of labor between the respondent as speaker and the researcher as listener. However, when sharing thoughts and emotions about personal experiences, participants want to feel that they are being understood and that what they say holds special meaning for the researcher. This requires the investigator to identify with the person being interviewed while remaining cognizant of his/her own feelings in order to focus on the responsibility of carrying out the research agenda. However, rather than occurring simultaneously, Bondi (2003) argues that this dynamic represents an oscillation between positions of immersion in the other's story and objective distancing by

the researcher. As she concludes, this oscillation "creates what might be described as room to maneuver, or as a kind of psychic space in which affinities and similarities can be recognized, at the same time as retaining a sense of difference and distance. Empathy can be thought of in terms of psychic space in which movement between positions is possible" (p. 73). Because life history research is intended to probe deeply into the experiential lives of others, empathy framed in this way can be a useful approach for conducting effective and meaningful interviews.

Managing and Analyzing the Data

The most significant task for the researcher in managing qualitative data is to transcribe the interviews and efficiently organize any supplemental material gathered in support of the study (Plummer, 2001). Effective management and organization of life history data is important because it facilitates meaningful and trustworthy analysis, interpretation, and reporting of findings. However, life history research generates a significant volume of information. For example, there can be as much as a one to ten ratio between the hours spent interviewing respondents and the hours needed to transcribe and analyze the data. This can make the act of transcription an arduous task. Among the strategies researchers can use to reduce time spent transcribing is to edit only those parts of the interview that are specifically relevant to the research topic (although this must be done carefully so that the broader analytic context is not lost) and to delete from the conversation speech hesitations, such as "uh" and "you know" (Plummer, 2001). Another option is to have the tapes edited by a professional transcription service. This can be expensive, but the advantage is that you can save a significant amount of time, which can then be devoted to data analysis and interpretation. Even if a professional service transcribes the tapes, researchers should listen repeatedly to each interview because it helps to identify possible editing errors, aids analysis by highlighting important ideas and themes, and facilitates intimate engagement with the respondent's stories. This latter point is especially critical "because intent and meaning are conveyed as much through how things are said as through the actual words that are used" (Goodson & Sikes, 2001, p. 33; for further insight, see Mishler, 1986).

A positivistic approach to research generally dictates that the act of analysis occurs only after all the data has been acquired. However, in qualitative research, and with life history studies in particular, the processes of gathering and analyzing evidence should be done simultaneously (Boyatzis, 1998). The goal for the researcher is to gain a better understanding of the co-constructed nature of the data as it emerges. The simultaneous gathering and analyzing of data also facilitates the exploration of possible new avenues of discovery with respondents. Cole and Knowles remind us that the researcher is always the primary instrument of analysis in life history

research and, therefore, "requires a kind of mental readiness to understand and accept the complexity of the task, the creative nature of the process, and the requirements of time, patience, and commitment to a sometimes convoluted and chaotic process" (2001, p. 99).

Although the practical act of qualitative data analysis can take many forms (Creswell, 1998), it often begins with coding the data into meaningful analytical units. Coding represents for the researcher the initial stage of interpreting how the respondent views the world and of constructing a story that draws the reader into the lived experiences of others. Whether coding life history data is done manually or with a qualitative analysis software program, such as NVivo or Atlas.ti, assigning codes involves reducing the text into categories that the researcher considers important in relation to the problems being studied (Spradley, 1979). Analysis can begin by either developing codes prior to examining current data, or the researcher may choose an inductive approach that allows codes to emerge as the text is being examined. Although either approach can be effective in allocating units of meaning to each respondent's story, I have found that the emergent nature of life history research generally supports an inductive approach to coding data.

The final analytical step is to arrange the codes into broad groupings that reflect general themes that inform a deeper understanding of the research problems being investigated. For example, in my life history project, statements made by respondents about key experiences in their lives that influenced their decision to enter librarianship (for example, working in the library as a college student) are coded and then arranged into broader conceptual categories (for example, pre-professional work experiences). Identification of conceptual categories helps the researcher determine where commonalities, differences, patterns, and structures of phenomena may exist. This creates the opportunity to raise possible new research questions, to show relationships across data, to delete or add codes, and to arrange codes into hierarchical order (Basit, 2003).

Presenting the Data

Qualitative research studies should always include a deliberate plan for how and in what form findings will be promulgated (Patton, 1987; Wolcott, 2001). In general, the presentation of life history findings can be framed in one of two ways. A study can present each life history as a specific case, followed by a summation of the critical issues that emerged from the analysis. This approach is helpful in highlighting the uniqueness of individual experiences. Another approach would be to present the findings thematically and supported by narrative excerpts drawn from interviews and other sources. Describing key themes that have emerged from the data can be effective in influencing policy because the experiences of respondents can be linked contextually to specific problems or assumptions.

Regardless of how the findings are framed, the goal of any qualitative report "is to bring understanding to complex social phenomena that cannot be reduced to precise, statistical relationships and . . . written in a style that uses literary sensibilities to take readers inside the issues and settings under investigation" (Cole & Knowles, 2001, p. 224). This style of writing raises important questions about how to present life history data effectively. For example, as the primary instrument of interpretation and analysis, the researcher should articulate any possible biases he or she may have about the study and its findings. Acknowledging possible biases reinforces the trustworthiness of the data and helps the reader understand the overall interpretive process used to examine the data. Another important issue is that the researcher must know who the intended audience is and what issues or decisions the study is intended to influence (Plummer, 2001). Knowing the intended audience is important for the librarian as a practitioner-researcher because the goal of a life history study in this context is most likely to provide evidence that could inform new ways of evaluating current practice and to document the uniqueness of individual users of library services and programs.

Life histories are an effective method for giving voice to those who may not otherwise be visible through other forms of inquiry. However, the challenge in writing up life histories is to develop a co-constructed story that respects and highlights the voice of the respondents yet also involves the author in the text as the principle instrument of analysis and interpretation. In addition, raw data cannot inform practice if it rests outside of the larger interpretative contexts brought forth by the researcher. This requires the researcher to be selective in presenting narrative excerpts that contextualize the data. Emerson, Fretz, and Shaw (1995) suggest the following criteria for editing and presenting the data: *length* (long quotes are difficult to read); *relevance* (link the data to the purpose of the study); *readability* (excerpts should make syntactic sense and not disrupt the overall flow of the text); *comprehensibility* (assure that the underlying meaning of a statement can be understood); and *anonymity* (any information that could reveal the identity of a respondent must be excluded) (p. 187).

Methodological and Interpretative Considerations

Space constraints do not allow for a detailed examination of all possible ethical and methodological dilemmas that could arise during a life history research study. The writings of Cole and Knowles (2001), Plummer (2001), and Goodson and Sikes (2001) should be consulted for complete examinations of pertinent issues. However, there are two critical issues that deserve special attention because they are particularly relevant to the study of libraries and librarianship by practitioner-researchers. These are (1) negotiating organizational insiderness and (2) understanding concepts of validity and researcher positionality in the text. I will follow this by sum-

marizing the potential benefits and pitfalls of applying life histories to the study of library organizations and conclude with a discussion of the ways life history research and other qualitative methodologies can enrich and broaden our understanding of library practice.

Organizational Insiderness

In a previous study I examined the hidden ethical and methodological dilemmas of being an insider participant observer (Labaree, 2002). However, several key points from this study are worth exploring further in the context of the life history method applied to the study of libraries and librarianship. Insiderness in qualitative research refers, in general, to the study of one's own culture or organization. Concomitantly, the concept of outsiderness refers to the act of examining a culture or organization that is unfamiliar to the researcher. A review of the literature highlights at least four implied advantages to being an insider. First, insiderness has value because the researcher will be familiar with the organizational setting and its members. Second, insider status has value because the researcher and the informant will have likely shared common social and occupational experiences (Cerroni-Long, 1994; Kanuha, 2000). The assumption is that experiential commonalities can form the basis for building trust and developing a relationship that contributes to a deeper understanding of the phenomenon being investigated. Third, insiderness implies that the researcher has a greater understanding of how to interpret cultural work habits and practices and obtain key information that is available only to organizational members. And fourth, insiderness has value because it facilitates reflexivity. Introspective analysis based on insider knowledge can lead to the discovery of greater clarity of purpose for the researcher and a deeper understanding of the evolving research process.

Despite these implied advantages, insider status is situational and dependent upon the underlying objectives of the study and level of access to key informants. In short, total immersion into the lived experiences of others can never be fully achieved. The life history researcher must, therefore, continually negotiate with respondents to ensure that mutual trust, access, and clarity of purpose is maintained. Imbedded within this process of continual negotiation are several ethical and methodological dilemmas. For example, an outsider must spend significant time and energy devoted to gaining entry into the research setting. The situatedness of being an insider researcher diminishes the need to gain entry. However, trust and cooperation must still be negotiated because the familiar colleague is now repositioned as a principle investigator of the organization. The added responsibility of studying and interpreting one's own community is especially challenging because any false representations of a phenomena, either real or perceived, could lead to feelings of betrayal on the part of the participants. The outsider has an equal responsibility to avoid false realities, but

they will eventually exit the research setting and become distanced from the consequences of inaccurate representations of people and places.

Another hidden ethical dilemma of being an insider is the challenge of unintended positioning and disclosure of data. This refers to the researcher's status within the organization and how one's position within that organization may influence how others respond to your study. For example, in my life history study, several respondents assumed that I "knew what was going on" with regards to their own socialization experiences because I had been working in the library for a number of years. However, because I wanted to document the respondent's own particular perceptions of reality and because I needed to clarify my own understanding of key issues, I had to remind the respondents that my position as an insider was governed, as well as limited by, my position as a faculty member within the library.

A related issue is the dilemma of shared and significant relationships. An insider researcher may need to interview or otherwise interact with close friends and colleagues. This is not inherently bad. In fact, Coffey (1999) notes that the position of being both an insider and an outsider requires social interaction so that trusting relationships with respondents can develop and grow. However, the issue of shared or significant relationships between the researcher and the researched is complicated by being an insider because the insider brings more to the social setting than previous knowledge about people and events. The research process also requires a newly negotiated dynamic between the researcher and the respondent that ties the two individuals together not only on the basis of collegiality and friendship but also on the additional basis of being a key informant in a life history study. The possible ramifications of damaging this first relationship in pursuit of the second must be clearly understood by the researcher and study participants.

Finally, insiderness influences the notion of disengagement from the research process and raises ethical questions about the researcher's obligations to informants after the study is completed. The issue of disengagement receives relatively little attention in the literature. However, for the insider conducting a life history study, the act of disengagement is critical because the researcher remains in the community once he/she has finished the research. Similar to negotiating trust and cooperation at the initial stage of the research process, disengaging from the research also carries possible repercussions from those who may feel that the study's conclusions or final recommendations are inaccurate. This requires the researcher to consider carefully who participates in the study and to measure how potentially sensitive findings are going to be reported and to whom. A good researcher will make certain that a study's purpose, goals, and potential impact on future decision making are understood thoroughly by everyone before, during, and after the research process.

Validity and Researcher Positionality within the Text

Validity and the author's place within the text are important to any life history study, whether it is being pursued for personal intellectual enrichment or it represents a study intended to evaluate policy or challenge existing assumptions about user behavior. In general, validity refers to the perception that findings are congruent with reality and that what the researcher is intending to investigate is really there (Merriam, 1998). Researcher positionality is concerned with the degree to which the author's place within the text is revealed; it relates to what Tierney describes as the "particular issues that we all face as we translate ourselves from researchers to writers" (1998, p. 52). Understanding the ethical and methodological constructs of validity and researcher positionality in life history research contributes to the reliability and trustworthiness of the findings and helps the reader determine whether there is a strong correlation between the author's interpretation of the data and any recommendations or conclusions presented.

For a life history study to provide insights into the experienced lives of others and to challenge successfully assumptions about current practice, the consumer of life history research must have confidence that the investigator has represented a valid reality of events and people. The challenge, of course, is that reality is subjective, multidimensional, and ever changing. It is, therefore, important to understand that the underlying purpose is not to describe "a reality" but to observe and document an individual's construction of reality (Merriam, 1998). Within this framework, the concept of validity relates to the confidence one has that the mode of analysis is actually investigating what it is supposed to investigate. As Plummer notes "If the subjective story is what the researcher is after, [then] the life history approach becomes the most valid method" (2001, p. 155). The challenge for the researcher in building a case for validity is to minimize perceptions of bias. The goal is not to achieve pure objectivity but to acknowledge and describe potential biases in a way that allows the reader to determine how these biases might influence their own interpretation of the findings. For example, the fact that I am a male researching the socialization experiences of several respondents who are female should be acknowledged because a consumer of the findings from my life history study might believe this to be important in determining how to ultimately interpret the data.

Bias can never be totally purged from a qualitative research study and, as Plummer (2001) notes, bias can arise from the respondent, from the researcher, or as a result of the interaction between the researcher and the respondent. Plummer suggests several validity checks that can be utilized to increase confidence in the findings. For example, the researcher can have the respondent read and critically examine all of the data from the study. This allows the respondent to reflect upon what has been said and to offer additional insight that may further contextualize initial findings. The researcher can also compare the life history data with other types of

biographical sources. As noted earlier, consulting secondary documents can help confirm factual information and determine the chronology of events. Finally, the researcher can strengthen validity by gathering data from individuals who may have knowledge about similar situations. This can include, for example, conducting brief, informal interviews with people who have similar backgrounds and experiences so that critical events described by the study's primary informants can be independently confirmed.

Closely related to the concept of validity in life history research is the issue of how the author presents him- or herself in the text (Tierney, 1998). As noted earlier, a challenge in reporting life history research is to develop a co-constructed story that illuminates the voice of the respondents while also acknowledging the author's role as the principle instrument of analysis and interpretation. Much of this discussion in the field of qualitative inquiry is wrapped up in the ambiguities of postmodernism (Prain, 1997), but here I will focus more on the practical problems of style and narrative voice.

Tierney (1997) argues that qualitative researchers generally present themselves in one of three ways within the text: as narrator, as interviewer, and as participant. As narrator, the author uses the "I" in the text to present a single narration of people and events (for example, "I interviewed Roscoe early Tuesday morning so that there was little chance of being interrupted"). The researcher can also take the position of interviewer. The dialogue is in the form of a question and answer exchange between the respondent and the researcher. The excerpt at the beginning of this article represents this type of textual dialogue. Finally, the author often enters the text as a participant, not only to help move the story along but to reveal a "human side to the discourse" (Tierney, 1997, p. 27).

The challenge underlying these imbedded textual identities is that they imply a stable narrator who simply enters the field, gathers and analyzes the data, and then reports the findings. The author's role in constructing reality is revealed as unproblematic and is often expunged from the text. The co-constructed nature of life history research requires a more deliberate strategy of not only revealing the voice of the respondent but of acknowledging and accepting a second voice in the text—that of the author. If consumers of life history research are to develop meaning and form judgments based on an interpretation of the text as it is viewed through their own realities, then the researcher must problematize "the privileged authorial perspective" (Webster, 1983, p. 195) most commonly found in social science research. Life histories demand a higher degree of authorial representation in the text because, at a fundamental methodological level, it is a journey of discovery between two individuals, the researcher and the respondent. It cannot be my story of his or her story, but rather *our* story revealed as a way to challenge existing assumptions and to document the interactions between experienced lives and the institutional and societal contexts in which they are lived.

Conclusion

Kouritzin (2000) identified several potential benefits of using life histories in organizational research. Applied to the study of libraries and librarianship, these benefits may include the following:

1. *Revealing the mundane.* Life histories allow perceptions about ordinary lives to become less ordinary. Life histories can be collected from individuals whose stories have not been documented before and, as a consequence, were never included in prior assessments of services and programs.
2. *Informing theoretical assumptions.* Life history research not only enriches general understanding but provides singular examples of experienced lives that may not fit within assumed theories concerning the relationship between people and organizations.
3. *Reinterpreting new knowledge.* Life histories are comprehensive and detailed. This feature means that the data about the lives of respondents can be revisited and reinterpreted as new knowledge or new theories are discovered.
4. *Developing contextual clarity.* Life histories are contextualized and historically grounded. This allows the reader to interpret policy decisions in human terms instead of framed only within economic, legal, or other research terms.
5. *Enhancing subjective awareness of others.* Life histories possess a literary and rhetorical style that makes them accessible to a wide audience and, as such, they help facilitate better understanding about "the untidy complexity of human decision making" (Gmelch, 1992, p. 38).

Life history as a research methodology also benefits the respondent and the researcher. For the respondent, the research process creates an opportunity to be listened to, perhaps for the first time, and a means for understanding and recognizing moments of experiencing adversity. Life history studies benefit the researcher, according to Kouritzin (2000), because they force the researcher to try to understand and then represent an emic perspective of social constructs. Life history research also creates opportunities to illuminate shared understandings about critical issues and events.

The research question underlying a study should always govern the choice of method used for analysis. Although life history research challenges conventional notions of what may be considered useful knowledge in assessing libraries and library practice, it requires a significant commitment of time on the part of both the researcher and the respondent. In addition, because the intention of a life history study is to develop a detailed profile of the experienced lives of others, a large volume of data is generated from the research process. Synthesizing this data into a cogent set of recommendations or guidelines for best practice requires a significant commitment by the researcher of both time and resources. Finally, it should be noted that,

for the academic librarian acting as practitioner-researcher, working with a college or university institutional review board (IRB) can be intimidating. The purpose of these boards is to ensure that the research complies with various federal laws intended to protect human participants from harm. The potential challenge of presenting qualitative research proposals before an IRB have been well documented (Lincoln & Tierney, 2004) and indicates the possibility that life history researchers must be prepared to justify in greater detail their chosen method of research than others relying on more conventional approaches.

Despite these challenges, life history research studies can reveal important new ways of understanding the relationships between libraries and the people who use them. Although there is recognition within the profession that applying qualitative research methods to the study of libraries and their users has value (see, for example, Bates, 2004; Dewdney & Harris, 1992; Fidel, 1993; Sutton, 1993; Westbrook, 1994), creative use of qualitative methods is not significant compared to that found in other applied social science disciplines. However, qualitative research, with its emphasis on understanding complex, interrelated, and dynamic phenomena, is particularly relevant to investigating the contextual features of contemporary libraries and librarianship. An important value of using qualitative research methods is that it provides a way to incorporate meaning as well as measurement into the way we understand library organizations and user behavior. I am not arguing that the use of life histories and other forms of qualitative inquiry replace prevailing positivistic and so-called evidence-based research methodologies. To do so would only encourage dichotomous debates about qualitative versus quantitative research paradigms. These discussions have little value in helping to address the problem of linking research and practice. However, expanding the application of qualitative research methods to the study of problems in librarianship will help challenge accepted conceptualizations of what has been defined traditionally as evidence in professional practitioner-research.

References

Basit, T. N. (2003). Manual or electronic? The role of coding in qualitative data analysis. *Educational Research, 45*(2), 143–154.

Bates, J. (2004). Use of narrative interviewing in everyday information behavior research. *Library & Information Science Research, 26*(1), 15-28.

Bondi, L. (2003). Empathy and identification: Conceptual resources for feminist fieldwork. *ACME: An International E-Journal for Critical Geographers, 2*(1), 64–76.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development.* Thousand Oaks, CA: Sage Publications.

Cerroni-Long, E. L. (1994). Insider and outsider perspectives in the study of ethnicity: The case of the Italian-American. *Ethnic Forum, 14*(1), 121–138.

Coffey, A. (1999). *The ethnographic self: Fieldwork and the representation of identity.* Thousand Oaks, CA: Sage Publications.

Cole, A. L., & Knowles, J. G. (2001). *Lives in context: The art of life history research.* Walnut Creek, CA: AltaMira Press.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage Publications.
Darby, L. (2005). Abolishing stereotypes: Recruitment and retention of minorities in the library profession. *Rural Libraries, 25*(1), 7-17.
Dewdney, P., & Harris, R. M. (1992). Community information needs: The case of wife assault. *Library & Information Science Research, 14*(1), 5-29.
Emerson, R. M., Fretz, R. I., & Shaw, L. L. (1995). *Writing ethnographic fieldnotes*. Chicago: University of Chicago Press.
Fiumara, G. C. (1990). *The other side of language: A philosophy of listening*. New York: Routledge.
Fidel, R. (1993). Qualitative methods in information retrieval research. *Library & Information Science Research, 15*(3), 219-247.
Gmelch, S. (1992). From beginning to end: An Irish life history. *Journal of Narrative and Life History, 2*(1), 29–38.
Goodson, I. (2001). The story of life history: Origins of the life history method in sociology. *Identity: An International Journal of Theory and Research, 1*(2), 129–42.
Goodson, I., & Sikes, P. (2001). *Life history research in educational settings: Learning from lives*. Philadelphia, PA: Open University Press.
Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. New York: Teachers College Press.
Harralson, D. M. (2001). Recruitment in academic libraries: Library literature in the 90s. *College & Undergraduate Libraries, 8*(1), 37-68.
Honma, T. (2005). Trippin' over the color line: The invisibility of race in library and information studies. *Interactions: UCLA Journal of Education and Information Studies, 1*(2). Retrieved February 12, 2006, from http://repositories.cdlib.org/gseis/interactions/vol1/iss2/art2.
Jarvinen, M. (2004). Life histories and the perspective of the present. *Narrative Inquiry, 14*(1), 45–68.
Johnson, J. M. (2002). In-depth interviewing. In J. F. Gubrium & J. A. Holstein (Eds.), *Handbook of interview research: Context and method* (pp. 103–119). Thousand Oaks, CA: Sage Publications.
Kanuha, V. K. (2000). "Being" native versus "going native": Conducting social work research as an insider. *Social Work, 45*(5), 439–447.
Koschmann, T., Zemel, A., Conlee-Stevens, M., Young, N., Robbs, J., & Barnhart, A. (2005). How do people learn? Members' methods and communicative mediation. In R. Bromme, F. W. Hesse, & H. Spada (Eds.), *Barriers and biases in computer-mediated knowledge communication: And how they may be overcome* (pp. 265–294). New York: Springer.
Kouritzin, S. G. (2000). Bringing life to research: Life history research and ESL. *TESL Canada Journal, 17*(2), 1–35.
Kvale, S. (1996). *InterViews: An introduction to qualitative research interviewing*. Thousand Oaks, CA: Sage Publications.
Labaree, R. V. (2002). The risk of "going observationalist": Negotiating the hidden dilemmas of being an insider participant observer. *Qualitative Research, 2*(1), 97–122.
Lincoln, Y. S., & Tierney, W. G. (2004). Qualitative research and institutional review boards. *Qualitative Inquiry, 10*(2), 219–234.
Merriam, S. B. (1998). *Qualitative research and case study applications in education* (2nd ed.). San Francisco, CA: Jossey-Bass.
Mishler, E. G. (1986). *Research interviewing: Context and narrative*. Cambridge, MA: Harvard University Press.
Morse, J. M. (1994). Designing funded qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 220–235). London: Sage Publications.
Patton, M. Q. (1987). *How to use qualitative methods in evaluation*. Newbury Park, CA: Sage Publications.
Plummer, K. (2001). *Document of life 2: An invitation to a critical humanism*. Thousand Oaks, CA: Sage Publications.
Prain, V. (1997). Textualizing your self in research: Some current challenges. *Curriculum Studies, 29*(1), 71–85.
Pritchard, I. A. (2002). Travelers and trolls: Practitioner research and institutional review boards. *Educational Researcher, 31*(3), 3–13.

Richie, B. E. (2001). Challenges incarcerated women face as they return to their communities: Findings from life history interviews. *Crime and Delinquency, 47*(3), 368-389.
Roberts, B. (2002). *Biographical research.* Philadelphia, PA: Open University Press.
Rubin, H. J., & Rubin, I. S. (1995). *Qualitative interviewing: The art of hearing data.* Thousand Oaks, CA: Sage Publications.
Sawyer, P. R. (2005). *Socialization to civil society: A life history study of community leaders.* Albany, NY: State University Press of New York.
Seidman, I. (1998). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (2nd ed.). New York: Teachers College, Columbia University.
Shotter, J. (1993). *Conversational realities: Constructing life through language.* Thousand Oaks, CA: Sage Publications.
Spradley, J. P. (1979). *The ethnographic interview.* New York: Holt, Rinehart and Winston.
Sutton, B. (1993). The rationale for qualitative research: A review of principles and theoretical foundations. *Library Quarterly, 63*(4), 411-430.
Tierney, W. G. (1998). Life history's history: Subjects foretold. *Qualitative Inquiry, 4*(1), 49–70.
Tierney, W. G. (1997). Lost in translation: Time and voice in qualitative research. In W. G. Tierney & Y. S. Lincoln (Eds.), *Representation and the text: Re-framing the narrative voice* (pp. 23–36). Albany: State University of New York Press.
Webster, S. (1983). Ethnography as storytelling. *Dialectical Anthropology, 8*(3), 185–205.
Westbrook, L. (1994). Qualitative research methods: A review of major stages, data analysis techniques, and quality controls. *Library & Information Science Research, 16*(3), 241-254.
Wolcott, H. F. (2001). *Writing up qualitative research* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Robert V. Labaree, Ed.D., is the Applied Social Sciences Librarian at the University of Southern California. He has been the subject specialist for international relations, political science, public administration, and urban planning since his arrival at USC in 1995. He received a doctorate in Higher Education Leadership from the University of Southern California, an M.L.S. from Louisiana State University in Baton Rouge, and a B.A. in History from the University of Nebraska–Lincoln. Recent publications include a dissertation study examining efficiencies of information exchange in academic decision making, a discursive analysis of organizational defense strategies during times of crisis, and an analysis of the ethical and methodological dilemmas of conducting insider research. Labaree is the principle investigator for three ongoing research projects: a five-year life history study that investigates the personal transition experiences of seven individuals who have made the decision to enter the field of librarianship; a study intended to develop a holistic understanding of the library as place based on the conceptual writings of geographer Donald Meinig; and a case study analysis that examines critically the challenges of promoting and sustaining effective communication throughout the design and implementation of a merger between academic libraries and computing service centers. He has served on the editorial board of *RUSQ* since 2002 and participated in several committees within ACRL and RUSA.

Identifying Opinion Leaders and Elites: A Longitudinal Design

Susan E. Davis

Abstract

The term *longitudinal design* refers to a flexible research approach that can be applied to a wide range of topics involving change over time. Longitudinal refers to both the data collected and the methods of analysis used, and project designs can combine several data-gathering and analysis methods within a longitudinal framework. Longitudinal research demonstrates several features that permit the observation of process and change and facilitate identification and evaluation of the underlying factors. Several library and information science studies demonstrate the application of a longitudinal approach to both prospective and retrospective research questions. This article draws primarily on a longitudinal study of leaders who emerged in the archival profession during the 1980s when archivists developed the first set of descriptive standards (MARC AMC) in response to trends in the automation of library cataloging. The study identified a core group of leaders whose influence drove the archival profession to move in a specific direction. The identification of opinion leaders and elites, and the factors that led to their status, has significant implications for understanding patterns of decision making and communication within organizations.

Introduction

The terms *longitudinal design* and *longitudinal analysis* apply to a wide range of research studies conducted within many social science disciplines. The concept of longitudinal research relates to both the nature of the data and the methods of analysis. Because researchers can use a longitudinal approach in combination with other methods, as well as by itself, the lon-

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 140–157
© 2006 The Board of Trustees, University of Illinois

gitudinal nature of a study is not always obvious. Research that focuses on process, adaptation, or change is often longitudinal, whether or not that element of the design is clearly articulated. The common denominator in all cases is time; in longitudinal research a span of time provides the crucial insight into the questions being studied.

This article examines longitudinal design and analysis as a research method, reviewing ways in which researchers have defined and applied this approach. Examples of longitudinal studies in library and information science (LIS), broadly defined, provide context for understanding why researchers choose this method and its strengths and weaknesses. These examples illustrate the kind of problems for which a longitudinal design is appropriate.

The article also focuses on a longitudinal research study by the author that analyzed changes in the archival profession during the 1980s, a period that experienced rapid change within archives and the adoption of the first set of descriptive standards, that is, the US MARC format for Archival and Manuscript Control (MARC AMC) (Davis, 2003b). The study sought to identify the opinion leaders within the profession and to understand how they were able to persuade and mobilize archivists to undertake what was viewed as a radical change in outlook and practice.

Definitions and Characteristics

The terms *longitudinal design* and *longitudinal analysis* have been used to describe a wide range of research using many different approaches. In fact, one of the strengths of longitudinal design is that it can serve as a framework for research that combines a longitudinal approach with other methods of data collection and analysis. Menard says that "longitudinal research must be defined in terms of both the data and the methods of analysis used in the research" (2002, p. 2). Following that line of reasoning, the term describes not one but a variety of methods that demonstrate the following features:

- Research covers a span of time in order to document process or identify change
- The direction of the research can be prospective or retrospective
- Data can be qualitative or quantitative
- Data should encompass multiple units of analysis
- Data collection should occur more than once

Researchers have elaborated on these points in their discussion of the method and its potential application to different kinds of research questions. Kimberly says: "Longitudinal organizational research consists of those techniques, methodologies and activities which permit the observation, description and/or classification of organizational phenomena in such a way that processes can be identified and empirically documented" (1976,

p. 329). He goes on to point out that the researcher defines the length of time for the project, as well as the research objectives, number of data collection periods, duration of time between collection periods, method of analysis, and unit of analysis. Venkatesh and Vitalari, who applied longitudinal analysis to information systems research, stated that "longitudinal research examines the behavior of processes and change in critical variables over time" (1991, p. 2). They also point out the benefits of using multiple methods to collect data in a variety of forms. According to Janson, "a 'longitudinal study' can be any diachronous study or a study of a process of change" (1981, p. 20). *Diachronous* refers to looking at changes over time, in contrast with *synchronous,* which is the analysis of factors existing or arising at a single point in time. This time span aspect is the primary factor that distinguishes longitudinal research from other approaches.

The basic definition does not specify whether the chronological direction is forward or backward. The majority of longitudinal research is prospective, however, because it is easier to plan to collect specific information in the future than it is to derive it after the fact. As King stated:

> The main difference between prospective and retrospective designs is the length of the recall period. In prospective designs the recall is generally closer to and captured as the phenomenon unfolds, while retrospective designs require the participants to recall events that have happened in the past. Intervening experiences and events can interfere with the accuracy of data in recalled events. (2001, p. 10)

However, the risks are lessened through careful and probing questions from the researcher. In addition, not all retrospective data result from interviews; one can also tabulate data from secondary sources.

The definitions also do not specify whether the data collected and analyzed are quantitative or qualitative. Longitudinal research is frequently quantitative in nature, although it can combine both qualitative and quantitative approaches. Ruspini (1999) suggests that research based on longitudinal data can build bridges between qualitative and quantitative research traditions. The terms *qualitative* and *quantitative* can apply to both the data and the analytic techniques.

Whatever data is collected should encompass a number of units of analysis, and the data should be collected on those units at more than one point in the study to allow comparison over time. Janson (1981) limits longitudinal research to studies that collect data on the same individuals or units at multiple points and that also use data on several of those units. He suggests that longitudinal analysis is sometimes treated as synonymous with cohort analysis, where the term *cohort* "has a very wide meaning of any subpopulation of individuals (or other units) with a common characteristic" (p. 21). Researchers also contrast longitudinal research with cross-sectional research where measurement occurs only once for each subject or variable.

Within those definitions, research design and data collection can include a range of methods, including panel and cross-sectional designs, interviews, and survey research. Longitudinal studies have been undertaken within the fields of anthropology, community studies, education, psychology, health, and criminology and can focus on transitions, changes, and adaptations, as well as the impact of events and circumstances (Holland & Thomson, 2004). Longitudinal research has many advantages and is particularly appropriate for studying social change and the diffusion of innovations. Menard suggests that "For many, longitudinal research is touted as a panacea for establishing temporal order, measuring change, and making stronger causal interpretations" (2002, p. 1).

Researchers frequently examine organizations from the perspective of process. In doing so, they look at the interaction among variables, but that interaction is not static. Important questions and issues evolve dynamically, making it important to be able to assess the same variables at different points in time as well as any cumulative effects on those variables. Ruspini (1999) emphasizes longitudinal design's heuristic potential because the data allow analysis based on duration, as well as the measurement of differences among variables across time.

Longitudinal studies frequently use historical sources, especially when the data are gathered restrospectively. Longitudinal design, however, differs from historical methodology in the ways in which the data are analyzed. Historical research involves the description and analysis of past events designed to reconstruct and understand how and why those events occurred and the roles of various players in those events. Historical research does not require that the data cover a span of time or that consistent variables be identified and measured more than once.

As attractive and flexible as longitudinal design may appear, it is not the answer to all research questions, even those looking at change over time. Identifying multiple, consistent units of analysis, for which retrospective analysis at specified points in time can occur, is not always possible. Research requires a framework of developing action that allows for the segmentation of activity for analysis. Prospective designs are frequently elaborate and costly, and there is always the risk of attrition among research participants, which can call the project's conclusions into question.

Research Studies and LIS

Researchers in the library and information fields have utilized a longitudinal approach. Below are examples of some studies that illustrate the diversity of possible study designs, data collection, and data analysis. In each of these cases, the authors have labeled their studies as longitudinal. Two of the study designs collected data prospectively, one using a qualitative and one a quantitative approach. Three of the studies focused retrospectively, again with a mixture of quantitative and qualitative data and analysis.

Prospective Designs

Preece, Schoberth, and Heinzl (2003) looked at changes in the activities of online communities over time, with the goal of identifying enabling and inhibiting factors. Their first step was to develop a conceptual framework with which they could describe the development of communications activity, and they tracked 33,000 participants over a three-year period—a prospective approach. To conduct their longitudinal analysis they divided the period of observation into equidistant intervals and used quantitative methods to analyze their data. Their units of analysis were the individual users, and they determined the time periods that organized their data collection and analysis.

In the late 1980s and early 1990s, Kuhlthau undertook a series of studies investigating the search process of high school and college students. She developed a six-stage model of the search process and sought to understand how users moved through that process in the course of their work (see, for example, Kuhlthau, 1991). The stages she identified constituted the time frame for the repeated data collection. The design was prospective in that she mapped student progress through the stages using a combination of interviews, questionnaires, process surveys, and flow charts and looked at both demographic and cognitive factors. Kuhlthau's articles documenting her research are widely cited and serve as a basis for many other studies on information-seeking behavior. The longitudinal aspect of the research, based on qualitative data gathered prospectively, made the data particularly rich.

Retrospective Designs

Julien and Duggan (2000) used qualitative and quantitative analyses to assess the literature on information needs and uses. Their goal was to examine the development of research in this area of LIS over time, and they identified two time periods (1984–1989 and 1995–1998) as a basis for data collection and analysis. They also compared their findings to research already done for the years between their two defined time spans. Their variables included degree of interdisciplinarity evident in references cited and whether the research was concerned with users' cognitive processes and systems design use. The authors also identified the research methods used. Their study identified longitudinal trends, including the indication that such literature was increasingly appearing in scholarly versus professional journals.

Wang and colleagues (White & Wang, 1997; Wang & Soergel, 1998; Wang & White, 1999) studied changes in relevance judgments over time, defined as the duration of a research project. They focused on judgments of usefulness of literature in a preliminary literature search, during the project, and at the writing for publication/citing stage. They gathered data synchronously at the literature search stage and retrospectively about

three years later to gain information about the use and citation stages. The projects had been completed by the time of the second interview. They asked similar questions at each stage about relevance judgments and the factors affecting them and were able to follow bibliographic items judged relevant initially through subsequent stages. This research is a good example of a project that was not intended as a longitudinal one, but, by building on the original research, the authors were able to compare judgments and factors at different stages.

Mustonen-Ollila and Lyytinen (2004) analyzed three organizations' adoption of information system process innovations (ISPI) over a period of four decades. They describe their work as a qualitative case study using a longitudinal vertical research design. They looked at a retrospective sample from over 200 ISPI adoptions and identified significant differences based on computing era, type of innovation, and organization. They divided computing into four distinct eras (early computing, 1954–1965; mainframe, 1965–1983; office computing, 1983–1991; and distributed applications, 1991–1997) and distinguished among four types of ISPIs (baseline technologies, tools, description methods, and managerial innovations). Through a combination of semi-structured interviews and archival data, Mustonen-Ollila and Lyytinen (2004) found that many adoptions were outcomes of internal learning, more ISPIs occurred during times of prosperity, and most innovations took place at the project management level.

Identification of Opinion Leaders/Elites

The remainder of this article will describe my study in which I used a longitudinal design to explore the development of the archival profession during the 1980s. Specifically, the study was designed to identify the opinion leaders who led the activity, acting independently and/or on behalf of their employing institutions and professional associations. The study focused on the development and adoption of the first set of descriptive standards (MARC AMC) as the single critical event that most changed the archival profession. Description occurs in all repositories, regardless of size or sponsorship, and it represents a fundamental archival function in that it demonstrates the way in which archivists connect resources and researchers. Thus, changes in description and the implementation of standards dramatically affect the work of archivists in any institutional setting.

The study was based on the assumption that every profession wants to enhance its status and jurisdictional control and that effective leadership is a factor in achieving that goal. The study of a particular profession provides an opportunity to examine more closely the interaction of individuals and groups and the exercise of influence and power in advancing agendas over time. While the research evaluated influence on the part of both individuals and organizations, the emphasis lay more in identifying individual leaders who drove the activities that spurred the changes. Kadushin (1968) suggests

that decision making is a way to identify elites and leaders, beyond position and reputation. Thus, the research questions called for an approach that would identify the individual opinion leaders and elites who influenced the rest of the profession.

Elites is a term used by sociologists, among others, to designate individuals within larger groups who possess characteristics that set them apart, usually implying a level above the masses. In Bottomores's classic work on the topic, he proposed that the term be applied to "functional, mainly occupational, groups that have high status (for whatever reason) in a society" (1966, p. 14). Researchers focus on elites and opinion leaders for many reasons, including efforts to understand trends and shifts within society and organizations. These issues are significant, as elites and opinion leaders have the ability to convince others to adopt innovations, to change course in an organization or association, and to undertake a wide range of tasks and activities. Their behavior often sets a standard for others to follow. For the purpose of this article, the terms *elites* and *opinion leaders* are used interchangeably. The identities of elite members of a group are not always obvious since they may not relate to formal positions within a formal hierarchy; opinion leaders may not be the elected officials or the titular heads of organizations. Instead, one must look at a variety of factors within an organizational history and structure to determine the variables that distinguish members of a group who assume such leadership roles.

The process of adopting descriptive standards took more than a decade; thus, the research questions were best served by a longitudinal approach that facilitated the analysis of change across time. During that period (1977–1990), individuals emerged who were responsible for leading a series of projects in three stages that resulted in dramatic changes for the archival profession. The longitudinal design of this study identified these elites by looking at a series of variables at specific points in the three chronological phases of activity.

The status of descriptive standards was clear at the outset of this activity (1977) and again at the end of the period under study (1990). But the developments occurred incrementally as individuals participated in grant projects, committees, and task forces. Because it was longitudinal, the study was able to show both developing activity and changing relationships among individuals. While the data gathered were largely qualitative, quantitative methods were used to validate the conclusions reached through the qualitative analysis. This method for identifying the emergence of elites over a period of years has significant implications for the discussion of decision making, communication patterns, and diffusion of ideas within an organization, institution, or profession.

The analysis in this study focused on three different units of social organization within the chronological framework. First, the study looked at the archival profession as a whole in terms of its structure and dynamics

and the trajectory of events resulting in description standards. What transpired during the time period relating to the specific issue of descriptive standards affected the profession at large. This broad outline provided context for the events that occurred and the array of actors who participated in those events. The second level of analysis revolved around the series of groups that undertook descriptive standards work and constituted a more chronological analysis of what took place. These bodies evolved over time in response to pressures and opportunities from workplace and professional organizations. Library consortia and granting agencies supported the work of these groups. The third level of analysis, on which this article will focus, concentrated on the individuals who populated these groups. Who were these individuals, what roles did they play, and how—and when—did they emerge as significant players and opinion leaders?

Individuals played extremely influential roles in the process of descriptive standards development. An archivist may have become involved initially because of workplace responsibilities but then developed a reputation based on participation in task forces, committees, and working groups. Individual archivists presented papers, taught workshops, and sought and received appointments to positions of influence. Their work reflected personally upon them but also enhanced the reputation of their institutions. The reverse may also have been a factor; prestigious institutions with greater resources may have provided more opportunities for their employees. In either case, over time, certain individuals rose to prominence.

The research design for this study had two parts. While the discussion will emphasize the second, it is important to understand how the first part created the framework for the second. The first half of the study provided the background and historical context for the activities that took place between 1977 and 1990. This portion of the study established the longitudinal framework, identified the phases of activity, and used traditional historical and archival sources to paint a chronological picture that set the stage for the work of individuals (Davis, 2003a). The second half of the study followed a more sociological approach, examining patterns of relationships and concentrating on the individuals involved in the development of descriptive standards during these phases. The goal was to identify the elites, understand how they rose to positions of influence, and examine the relationships among these opinion leaders.

The archival universe during this period was comprised of individuals and organizations, operating both independently and in groups. For example, individual archivists worked in archival institutions ranging from colleges and universities, public libraries and museums, and government agencies in the federal, state, and local levels, to corporations, not-for-profit organizations, and religious associations. While archival principles (including description) remain fairly constant across those categories, the ways in which they are carried out vary according to the nature and size of the archival unit.

These individual archivists operated within the confines of their workplace, and that workplace may have defined both their priorities and the scope of their activities. In the case of descriptive standards, the existence of an online public access catalog (OPAC) within the larger institutions, such as university archives, often provided an incentive for the archivist to become involved in the area of descriptive standards for archival holdings.

Individual archivists were also involved in professional associations on the national and regional levels. While they participated as representatives of their employing institutions, they also participated because of their own professional interest and ambition. As self-identified members of a profession, many archivists have been active in advocacy for the profession. They have also worked toward improvement of practice, promotion of education, and in this case, the development and adoption of descriptive standards.

As is the case with longitudinal research, data collection on individuals fell into several stages, each of which will be described. First, it was important for me to understand the sequence of events that took place in order to identify the relevant groups and activities before I could pinpoint the individuals involved in each activity and the level of involvement. Once I accumulated the names of all the individuals, I then had to design a mechanism for differentiating levels of involvement so that the elites would begin to emerge. Those individuals became the subject of semi-structured interviews, conducted via phone or in person. Participants had the opportunity to name other influential individuals, and that data were tabulated and additional names added to the interview list. Data from the interviews, combined with data from the archival sources consulted in the early stages of the project, revealed relationships among the opinion leaders, and those were analyzed further.

Primary and secondary sources, consulted during the first part of the study, revealed that the events leading to the development and adoption of descriptive standards fell into three distinct chronological phases.[1] Prior to this period, archivists viewed description as the production of *finding aids,* or narrative descriptions of the creators, scope, and content of collections, accompanied by box and folder lists. Card catalogs frequently contained summary descriptions of archival collections, and archivists expected researchers to locate the desired material in the catalog, move to the longer register or inventory—as those narrative finding aids were often called—and then request specific containers of material. Archivists assumed that, because each collection was unique, standardized descriptions were impossible and that researchers could just continue to find their way to the repositories holding relevant resources.

The advent of computer applications, specifically in the area of bibliographic control, began to filter into the archival community. Archives frequently exist within libraries, and thus archivists felt pressure to comply with online catalogs and the MARC format. At the same time, government

archives were beginning to experiment with a program called SPINDEX (Selective Permutation Indexing). The archival profession, therefore, was faced with a decision regarding the appropriate course for online access to archival collections.

In response, during the 1977 annual meeting, the Society of American Archivists (SAA) Council established the National Information Systems Task Force (NISTF) to study the problem of constructing a national information system for archives and manuscript collections. NISTF and its work constitute the first chronological stage of descriptive standards activity, lasting from 1977 to 1983.

NISTF's work had three goals: to provide intellectual access to archives and manuscript sources in American repositories; to establish a framework for "describing and improving access to archival resources"; and to facilitate the adoption of automated techniques (*SAA Newsletter,* May 1981, pp. 6–7). Basically, the SAA Council charged them with determining the best direction for future efforts. As part of their work, NISTF completed a data elements dictionary that, while never published, served as an extremely influential document. The data elements dictionary demonstrated that the units of information used by archivists were sufficiently consistent to support communication across different systems. But the Task Force also concluded that a single national information system was unlikely because of resource issues and the diversity of archival repositories. NISTF's charge expired in 1983, and a new SAA standing committee, the Committee on Archival Information Exchange, took over its ongoing responsibilities.

During this first stage, members of the library and archival community were working with the Library of Congress (LC) and others to develop MARC AMC. The Research Libraries Group (RLG) was a particularly active participant in this process because it saw the implementation of the format as crucial to the inclusion of special collections holdings into their Research Libraries Information Network (RLIN). The American Library Association approved MARC AMC in 1983, and the same year LC published Steven Hensen's first edition of *Archives, Personal Papers, and Manuscripts: A Cataloging Manual for Archival Repositories, Historical Societies, and Manuscript Libraries,* which interpreted the MARC format for archivists.

By end of the first stage, the direction of future activity was fairly clear. The National Archives was not going to take a leadership role in this endeavor; instead, pioneering efforts were more likely to emerge from the library community, which already had a huge stake in the MARC format. The archival profession now had a mechanism for descriptive standards work, and a group of leaders was beginning to emerge to spearhead the adoption of MARC AMC. The individual members of NISTF represented both their institutions and the larger profession. They brought to the table professional expertise gained through their positions and thus could see both the benefits and pitfalls of proposed descriptive systems. These in-

dividuals were also positioned to test the new descriptive formats in their repositories and share the results with their colleagues at home and in the larger professional arena. Archivists working in RLG libraries had the added benefit of RLG's intense interest in the MARC format, as well as its financial and political support.

The second phase of activity lasted from approximately 1984 to 1988, following the approval of MARC AMC. If NISTF served as the catalyst to define the problem and outline the desired approach to developing automated archival descriptive systems, RLG and SAA provided the means to disseminate that information to the profession at large. RLG established an Archives, Manuscripts, and Special Collections Task Force in 1983 to solicit broad-based participation in the automation of archival description through this new format. Members of this Task Force had participated in some of the NISTF discussions as well as those at LC, whose Joint Committee on Specialized Cataloging was working on revisions to AACR2. RLG's Task Force became a committee and continued their work until 1992. In fact, several NISTF members and the majority of the original members of SAA's new standing Committee on Archival Information Exchange were employed by RLG institutions.

The other major activity of the second time period was a series of workshops sponsored by SAA with funding from the National Endowment for the Humanities (NEH). SAA began offering MARC AMC workshops in February 1986, and by mid-1987 it had held seven workshops attended by 170 people representing over 140 repositories (*SAA Newsletter,* March 1987, p.7). The success of the workshops encouraged NEH to extend funding for a second two-year period as well as underwrite the revision of *Archives, Personal Papers, and Manuscripts,* published in 1989. SAA held these workshops at their annual meetings as well as venues around the country. Others contributed toward the wider dissemination and adoption of MARC AMC through articles or presentations at professional meetings. A conference held at the State Historical Society of Wisconsin in October 1984, funded by the National Historical Publications and Record Commission (NHPRC), resulted in two volumes (*MARC for Archives and Manuscripts: The AMC Format* [Sahli, 1985] and *MARC for Archives and Manuscripts: A Compendium of Practice* [Evans & Weber, 1985]) that were the first major attempts by the profession to disseminate widely the ways in which institutions applied the specific MARC fields to their own archives and manuscripts collections.

By the time the NEH funding ended, SAA had strengthened its position as a focal point for archival automation and descriptive activity, and RLG had positioned itself as a pioneer in online access to archives and special collections. Many of the individuals active in the first phase continued their work in descriptive standards as workshop leaders, authors, committee members and chairs, and presenters at meetings. They served as spokespeople for standards development and solidified their own reputations

as well as the work of the profession. Other leaders also emerged during this phase.

The third phase (1988–1990) consolidated the activities of the previous decade and provided the profession with a road map for addressing long-term standards development and implementation. SAA and RLG had engaged in a wide range of activities up to this point, but RLG had focused on its member institutions rather than a broad cross-section of the population. As a voluntary association, SAA lacked enforcement authority.

In 1988 a group of archivists representing a range of public and private institutions met to consider the larger questions of identifying and implementing standards. They successfully sought funding for their work from the NHPRC, and between 1988 and 1989 the Working Group on Standards for Archival Description met twice and drafted a new definition of description that incorporated its ongoing nature and focused more on the process than the end result of specific finding aids. In addition, the Working Group developed a matrix that articulated the levels of description, their relative strengths, and the sources of the various archival descriptive standards. Two issues of the *American Archivist* (Fall 1989 and Winter 1990) contain their final report and recommendations as well as the background papers and lists of additional resources.

It is significant that the Working Group emerged through the concern of individuals who, by and large, had become leaders in this area through the activities of the first two phases. Neither SAA, RLG, nor LC sponsored this project, although the members represented their interests. The Working Group constituted a group of opinion leaders within the archival profession who took the initiative and successfully grappled with the theoretical and practical issues of archival standards development and implementation.

This chronological saga is essential for understanding what transpired, for identifying the three points for data collection, and for identifying individual participants who played significant roles. Using the primary and secondary sources, I identified eighty-five individuals who had given relevant papers, written articles, taught workshops, or were members of NISTF or the Working Group. I designed an Excel spreadsheet, entered data on each individual's participation (NISTF member, number of conference presentations, etc.), and noted whether they worked for an RLG member institution. I also added data regarding other professional leadership positions, such as being named an SAA Fellow (the highest individual honor in the profession) or being elected an SAA Council member or officer.

The next step was to reduce eighty-five individuals to a manageable number for interviewing and later data analysis. I assigned weights to each category of participation, based on my assessment of the significance of that activity or honor, and applied those weights to each of the eighty-five names in the spreadsheet (Davis, 2003b). The weighting process reduced the eighty-five to twenty-three individuals who clearly had a higher degree

of participation than their colleagues (more than eight points). Two of those individuals had passed away, so the initial list of interviewees was twenty-one.

I contacted each of these individuals to introduce the project and request an interview. No one was reluctant to talk to me, and all agreed to have their interviews taped and then transcribed. I conducted a semi-structured interview, either in person or by telephone, with each of the twenty-one individuals who had risen to the top. Most interviews took at least an hour to complete. To minimize bias, I alternated the order in which I conducted the interviews to vary factors such as gender, affiliation with public or private repositories, and the chronological period in which the participant was active. I numbered the interviews based on the sequence in which I conducted them and used those numbers as identifiers throughout the analysis.

Each interviewee had the opportunity to name individuals he or she thought were influential in the descriptive standards process, including him- or herself (which a few did!) I designed a matrix and tabulated "chooser/chosen" data in order to confirm or amplify the names revealed in the written records. Tabulating the "chosens" resulted in an "average influence score" for each individual. This process served two purposes. First, it resulted in the addition of two more participants based on their number of influence points, bringing the total number of interviewees to twenty-three. The two individuals had missed the initial cutoff largely because they were not archivists and belonged to fewer of the relevant groups. Second, the process served as a validity check on the twenty-three individuals selected for interviews. I later also used the influence points in the data analysis.

In response to questions, the interviewees reflected on their own role in the descriptive standards process. In particular, I was interested in the reasons for their participation and whether the incentive came from their employer, their affiliation with SAA, or other personal ambitions. I also asked them to speculate on the effect this activity had on their career in terms of job opportunities, professional advancement, or personal growth.

Clear patterns of activity and participation emerged from the data collection process. These patterns served multiple purposes: to identify the elites and the reasons they rose to prominence, and to establish the stability of leadership and the relationships among the opinion leaders. The interviews amplified the information revealed through the written records. It was clear that certain individuals were more prominent in the early years, others became involved later in the 1980s, and others took active roles throughout the decade. The individual-level weighted data file clarified the roles and relative prominence of individual archivists and served as an objective way to determine the most likely leader candidates and those most appropriate to interview. The interviews provided information that complemented the other sources and validated the selection of interviewees. For data analysis, I added the two individuals who were deceased to the core group of elites for

a total of twenty-five units. Thus, the longitudinal analysis focused on these twenty-five opinion leaders at the three chronological stages of activity.

I chose to use network analysis as the framework for the data analysis. Network analysis depicts connections among individuals and clarifies who held positions of leadership and how that occurred. "A basic strength of the whole network approach is that it permits simultaneous views of the social system as a whole and of the parts that make up the system" (Wellman & Berkowitz, 1988, p. 26). The descriptive standards process was a complex process that involved individuals and organizations through a series of subgroups and projects. Network analysis is an appropriate way to explicate the roles, as well as strengths, of these individuals and organizations through their various activities.

In order to undertake the network analysis, I set up a series of matrices, using Excel, that detailed the connections among the twenty-five opinion leaders. Using the matrices, I was also able to compare data based on specific variables. The data from these matrices were imported into Ucinet, a social network analysis software that generates centrality measures (Borgatti, Everett, & Freeman, 2002). These centrality measures, displayed in various ways, documented the intensity of the relationships among individuals and how those relationships shifted according to context and time. I also calculated degrees of centrality across time periods using Pearson's product moment correlation and Spearman's rank order correlation.

The matrices were based on the fact that the development of descriptive standards fell into three fairly distinct chronological periods, differentiated by a series of task forces and grant projects that accomplished stages of the work. These issues brought people into the activities and established a specific set of connections that linked individuals for at least that span of time. Thus, the time periods represented the basic framework for the network analysis.

For each period I looked at three foundations of personal relationships: (1) participation in a specific issue or project (for example, NISTF); (2) a contextual set of relationships (for example, RLG affiliation); and (3) an element of personal connection (when an individual says he/she entered the profession).[2] I created a series of sociograms that visually documented the connections. This method of data analysis is consistent with longitudinal design in the focus on multiple units for whom similar information was collected at multiple points in time.

The connections among individuals based on each of these factors were symmetrical; everyone in each category was equally connected with everyone else in the category. It was also useful to look at the relationships across these content areas within each of the time periods. Because there were three foundations for personal relationships in each time period, each individual had the potential for one, two, or three connections with any other individual active in that time period. This involved "stacking"

the three sociograms for each time period, resulting in data that combined individual attributes for each phase, indicating strength of ties as well as the ongoing bases for connections among the elites.

These measures reflected individuals' positions over the duration of the study. I produced a similar matrix for influence, based on the chooser/chosen data collected during the interviews, imported that data into Ucinet, and correlated the information with the position matrices. Influence, as recorded in the chooser/chosen mentions, relates to an individual's reputation. Therefore, it was possible to examine the relationship between position (what an individual did) and reputation (the opinion of others) over time. In all of these instances, quantitative data derived from the qualitative sources facilitated the kind of consistent comparisons across time that longitudinal design demands.

Diffusion of innovations is another significant theoretical construct that underlies the analysis of the data. Rogers's (1995) landmark book on the subject describes the stages through which innovations are spread as well as the roles of individuals and organizations in the process. It is possible to examine the development of descriptive standards and place the phases along the continuum that Rogers delineates. It is also relevant to look at the adopter categories Rogers identifies, including the innovators and early adopters who represent the leaders this research sought to understand. Longitudinal design is an excellent tool for looking at the diffusion of innovations (see, for example, Mustonen-Ollila & Lyytinen's [2004] study).

The methods previously described revealed a cohort of twenty-five opinion leaders whose efforts drove the archival descriptive standards process. The names surfaced initially in the written records and were validated through use of the weighting scale and the interviews. The relationships among the opinion leaders and the reasons for and strengths of those relationships became clear through the social network analysis. Thus, findings relate to both the identification of the elites and understanding the structure of the interpersonal relationships and the way those relationships changed over time.

The analysis revealed certain characteristics that defined the opinion leaders—information that this research method made possible. First, members of the elite group were those who became active in the early or middle stage; most would be considered early adopters of the format. Second, factors such as association with an RLG library or other policy-making organization (for example, NHPRC) were more important for gaining leadership than being in a practice position within an institution. SAA membership was significant for credibility, but SAA membership alone did not lead to positions of influence. Demographics also surfaced as a significant factor. The individuals who comprised the leadership group largely belonged to a demographic cohort—those who came into the profession at the early stage of descriptive standards development and thus were in the right place at the right time.

Position did play a role in initial leadership status. Initially, one had to be in an institution ready to participate in this new venture. But beyond the initial stage, reputation took on increasing importance. Some individuals remained within their institutions; others changed jobs. But once an individual's name was connected to descriptive standards work, he/she could continue to participate, and further involvement rested on both reputation and continued interest. Indeed, several individuals rose to general positions of professional prominence as a result of their work on descriptive standards; many became SAA Fellows as a result of this work.

Members of this elite group exerted what French and Raven (1968) define as "expert power," a form of power based on the notion of special knowledge held by those with power. These opinion leaders were successful in codifying professional knowledge, and that knowledge became an essential element of professional practice. The leadership cohort represented not just an aggregate of individuals but also a cohesive group that drew strength from shared concerns and values. As Perrucci and Pilisuk found, "there exists in communities a relatively small and clearly identifiable group of *interorganizational leaders*" (1970, p. 1044). They concluded that such ties can "result in the creation of resource networks which can be mobilized and brought to bear upon particular community issues" (p. 1056). The pooling of resources that occurs in such a group enhances and expands individual power. The Working Group is a good example of such interorganizational leaders.

Conclusion

The approach taken for this research was somewhat inductive. The study began with a series of research aims, including (1) to identify individuals who played influential roles in the development of descriptive standards and how that changed over time, and (2) to analyze whether and how their organizational affiliations related to their influence in the descriptive standards process. The specific methods used for data collection and analysis evolved over the course of the project. The more historical methodology of the first part of the project explicated the settings and contexts for activity, and those findings drove the structure of the rest of the research.

The study is consistent with the characteristics of longitudinal design outlined at the beginning of this article. The research documented activity and change over a span of time, in this case approximately thirteen years. The time period fell into clearly defined phases, facilitating data collection at multiple points. Although largely qualitative in nature, the retrospective data were manipulated statistically and corroborated the qualitative analysis. Twenty-five individuals constituted the units of analysis emerging from the initial data collection. I collected data on those individuals at three points in time from a variety of sources and compared information across participants and across time periods according to several factors. The resulting analysis

pointed to trends that help us understand the process through which a profession evolves, develops standards, and codifies knowledge.

Longitudinal design and social network theory were appropriate frameworks for identifying elites. Historical methodology was crucial for building the context and understanding the course of activity. However, historical methods would not have allowed me to map the stages of activity and the shifting roles of the participants as well as perform longitudinal and social network analysis, which clarified roles and levels of influence. The array of individual attributes reflected specific aspects of individual's lives that contributed to their status as elites. The network analysis permitted comparison of positional elements to an individual's reputation. Thus, the application of social network theory to the overall longitudinal design allowed for the quantification of data that contributed to the validity of the findings.

This article demonstrates that flexibility is a major strength of longitudinal design. Each of the studies mentioned in this article illustrates ways in which researchers have pursued a range of research topics through studies that incorporated the elements of a longitudinal study with other methods of data collection and analysis. Each study established a chronological framework, facilitating data collection on multiple units of analysis more than once. Using prospective or retrospective approaches, quantitative or qualitative data, or a combination of all of the above, researchers have been able to draw conclusions regarding causality, organizational processes, and patterns of change. Longitudinal design is an excellent method that has been and should be applied to many settings.

Notes

1. These primary and secondary sources included published journal articles, the bimonthly newsletter of the Society of American Archivists (SAA) (titled *SAA Newsletter* during that time period), newsletters from the Research Libraries Group (RLG), archival records from the Society of American Archivists, records from the Bentley Fellowship Program at the University of Michigan and the National Historical Publications and Records Commission (NHPRC), an unpublished report from the National Archives and Records Administration (NARA), and personal files lent by colleagues documenting committee and task force work.
2. During the interview process, it became clear that many of the participants had entered the profession at the same time, in the 1970s.

References

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *Ucinet for Windows: Software for network analysis.* Harvard, MA: Analytic Technologies.

Bottomore, T. B. (1966). *Elites and society.* Middlesex, England: Penguin.

Davis, S. E. (2003a). Descriptive standards and the archival profession. *Cataloging & Classification Quarterly, 58*(3/4), 291–308.

Davis, S. E. (2003b). *Organization and influence in professional standards development: The case of archival description.* Unpublished doctoral dissertation, University of Wisconsin–Madison.

Evans, Max J., & Weber, Lisa B. (1985). *MARC for archives and manuscripts: A compendium of practice.* Madison,: State Historical Society of Wisconsin.

French, J. R. P., & Raven, B. (1968). The bases of social power. In D. Cartwright & A. Zander (Eds.), *Group dynamics: Research and theory* (pp. 259–269). New York: Harper & Row.

Hensen, S. (1983). *Archives, personal papers, and manuscripts: A cataloging manual for archival repositories, historical societies, and manuscript libraries.* Washington, DC: Library of Congress.

Holland, J., & Thomson, R. (2004). *ESRC study into the feasibility of a possible qualitative longitudinal study.* Retrieved November 2, 2005, from http://www.ccsr.ac.uk/methods/festival2004/programme/Thu/am/C/documents/ThomsonSummary_000.doc.

Janson, C. G. (1981). Some problems of longitudinal research in the social sciences. In F. Schulsinger, S. A. Mednick, & J. Knop (Eds.), *Longitudinal research: Methods and uses in behavioral science* (pp. 19–55). Boston: Martinus Nijhoff.

Julien, H., & Duggan, L. J. (2000). A longitudinal analysis of the information needs and uses literature. *Library & Information Science Research, 22*(3), 292–210.

Kadushin, C. (1968). Power, influence and social circles: A new methodology for studying opinion leaders. *American Sociological Review, 33*(5), 685–698.

Kimberly, J. R. (1976). Issues in the design of longitudinal organizational research. *Sociological Methods and Research, 4*(3), 321–347.

King, M. P. (2001, June). Cross-sectional and longitudinal research designs issues in the studies of human development. *Graduate Research in Nursing,* 3. Retrieved November 2, 2005, from http://www.graduateresearch.com/King.htm.

Kuhlthau, C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science, 42*(5), 361–371.

Menard, S. (2002). *Longitudinal research* (2nd ed). Thousand Oaks, CA: Sage.

Mustonen-Ollila, E., & Lyytinen, K. (2004). How organizations adopt information system process innovations: A longitudinal analysis. *European Journal of Information Systems, 13*(1), 35–51.

Perrucci, R. & Pilisuk, M. (1970). Leaders and ruling elites: The interorganizational bases of community power. *American Sociological Review, 35*(6), 1040–1057.

Preece, J., Schoberth, T., & Heinzl, A. (2003). Online communities: A longitudinal analysis of communication activities. *Proceedings of the Hawai'i International Conference on System Sciences, Jan 6–9, 2003.* Retrieved November 2, 2005, from http://www.ifsm.umbc.edu/~preece/paper/9%20HICSSNOCD06v2.pdf.

Rogers, E. M. (1995). *Diffusion of innovations.* New York: Free Press.

Ruspini, E. (1999). Longitudinal research and the analysis of social change. *Quality and Quantity, 33*(3), 219–227.

Sahli, Nancy. (1985). *MARC for archives and manuscripts: The AMC format.* Chicago: Society for American Archivists.

Venkatesh, A., & Vitalari, N. P. (1991). *Longitudinal surveys in information systems research.* Center for Research on Information Technology and Organizations (CRITO), UC-Irvine. Retrieved November 2, 2005, from http://www.crito.uci.edu/noah/NOAH/paper/LongSurveyInfoSys.pdf#search='longitudinal%20research%20retrospective; originally appeared in Kramer, K. (Ed.). (1991). *The information systems challenge: Survey research methods.* Cambridge, MA: Harvard University Press.

Wang, P., & Soergel, D. (1998). A cognitive model of document use during a research project, study I: Document selection. *Journal of the American Society for Information Science, 49*(2), 115–133.

Wang, P., & White, M. D. (1999). A cognitive model of document use during a research project, study II: Decisions at the reading and writing stages. *Journal of the American Society for Information Science, 50*(2), 98–114.

Wellman, B., & Berkowitz, S. D. (1988). *Social structures: A network approach.* Cambridge: Cambridge University Press.

White, M. D., & Wang, P. (1997). A qualitative study of citing behavior: Contributions, criteria, and metalevel documentation concerns. *Library Quarterly, 67,* 122–154.

Susan E. Davis, Ph.D., is an Assistant Professor at the University of Maryland, College of Information Studies, where she coordinates their specialization in Archives, Records, and Information Management. Her research focuses on archives administration and education as well as professional leadership and standards development. She has been an active member of the archival profession for many years, has taught a range of archival topics at several universities, and is a Fellow of the Society of American Archivists.

Meta-Analysis in Library and Information Science: Method, History, and Recommendations for Reporting Research

Matthew L. Saxton

Abstract

Meta-analysis is a method for summarizing statistical findings across multiple research studies. It is a useful method for assessing the level of agreement or disagreement surrounding a given research question. The ability to perform meta-analysis is dependent on the level of consistency in measures and the amount of data shared in published research. Guidelines to minimum standards for reporting research may improve the quality of writing in published research. Inconsistencies in reporting research findings across studies, failing to provide enough detail on method and instrumentation to facilitate replication, and the multiplicity of different operational definitions or measures for the same concept all pose difficulties to successfully attempting any form of research synthesis. This article presents a methodological explanation of meta-analysis, a literature review describing the application of meta-analysis in library and information science, and guidelines for reporting quantitative research that would enable subsequent researchers to perform meta-analysis.

Introduction

Every scholarly journal provides highly precise guidelines to its authors regarding the length of articles, the formatting of manuscripts, and the style of citations and footnotes. While authors may meet these guidelines with varying degrees of success, at least all parties involved in the scientific communication process recognize that a standard has been established. Curiously, few scholarly journals provide any guidelines regarding standards for the reporting of research in terms of the descriptive elements of

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 158–170
© 2006 The Board of Trustees, University of Illinois

a dataset that should be shared, the statistics that should be presented for a given method of analysis, and whether or not a copy of the instrument should be included. One reason for this omission in the field of library and information science (LIS) may be because of the variety of disciplinary and methodological approaches being used by researchers. To impose rules for the reporting of research might curtail the creative freedom of authors in presenting their work. However, this rich variety of quantitative and qualitative methods and different disciplinary orientations argues all the more for such guidelines to be established. For example, whereas physics or economics may have more rigid rules for publishing research that are well understood by researchers in their respective disciplines, LIS encompasses a much broader array of research methods that is harder to explicitly articulate. How does a researcher specializing in information retrieval working with a database of 10,000 records and hundreds of queries know how to evaluate a piece of research on information behavior based on twenty in-depth interviews? How does a researcher studying information services who reviews thousands of virtual reference transactions understand the validity of a philosophical investigation in classification theory? Such confusion may grow worse when LIS researchers examine the work of their colleagues in computer science, management, law, health informatics, or technical communications whose research questions may be similar to our own.

A guide to the minimum standards for reporting research may serve to help nonspecialists (as well as students) better understand what to expect when reading about a study employing a method with which they are unfamiliar. A second and perhaps more important benefit might be to improve the quality of writing in published research. Does the article provide enough detail so that the study could be replicated? Does the article then provide enough data so that results from a subsequent study could be compared to findings from the original study? Without replication, research in LIS advances haltingly, and validation of findings is difficult to achieve. The development of commonly accepted definitions and indicators for important concepts proceeds slowly. How do we measure information anxiety, collection strength, or user satisfaction? With the absence of a predominant method of observation, researchers often develop their own operational definitions for each new study. Even when discussing relatively concrete concepts such as number of volumes in the collection, different sources use different measures (compare the Association of Research Libraries [ARL] statistics to guidelines on counting given by various state libraries), and members of the ARL debate what it means to "own" volumes placed in a regional repository (ARL Committee on Statistics, 1997).

Inconsistencies in reporting research findings across studies, failing to provide enough detail on method and instrumentation to facilitate replication, and the multiplicity of different operational definitions or measures

for the same concept all pose difficulties to successfully attempting any form of meta-analysis. Meta-analysis is a form of research synthesis, and the terms are used interchangeably in fields that rely heavily on quantitative methods. Meta-analysis is a body of techniques that enables researchers to draw conclusions based on the findings of previous studies and present them in a useful and compact fashion (Matt & Cook, 1994; Hunter & Schmidt, 1990). The benefit of meta-analysis is that it enables researchers to obtain a greater understanding of the nature of the association between outcome and independent variables by comparing different values of effect size gathered from a large body of research. The ability to summarize findings across multiple situations and discover consistent trends (or in some cases, inconsistent trends) is a critical component of scientific research.

The lack of common definitions and research replication may be explained by two factors. In terms of number of researchers, number of Ph.D. graduates, and amount of available research funding, LIS is clearly a much "smaller" field in comparison to the sciences and other social sciences. Also, the field has a growing number of new scholars as many graduate schools expanded their doctoral programs from 1995 to 2005 in response to a growing awareness of the looming shortage of new faculty. Original research and the introduction of new methods enables younger faculty to build a stronger case for tenure (ironically, the author's own interest in meta-analysis is just such an example of this behavior). Nonetheless, maturity of a research area cannot be achieved without consensus building among scholars, repetition of studies or experiments to validate findings, and research articles or books that represent what Boyer (1990) defines as the scholarship of synthesis. Meta-analysis is a useful methodology for assessing the level of agreement or disagreement surrounding a given research question, and the growth in the number of meta-analytic studies in the literature is itself an indicator of increasing maturity in a given research area.

This article begins with a brief methodological explanation of meta-analysis and refers the reader to further sources for information on how to perform this type of study. This is followed by a literature review explaining the application of meta-analysis in library and information science or closely related fields. In conclusion, the author presents a set of guidelines for reporting quantitative research that would enable subsequent researchers to perform meta-analysis (and also increase the likelihood of having one's own research included in such subsequent study).

Meta-Analysis: Nuts and Bolts

Bivariate analysis involves examination of the extent to which one variable may have an influence on another variable, often described as the ability of one variable to *predict* (but not necessarily cause) the value of the other. Correlation and cross-tabulation are two common forms of bivariate analysis. *Effect size* is a measure of how much change in the dependent variable

can be predicted by the independent variable. A correlation coefficient is a common form of estimating effect size. The overall process is relatively straightforward and easy to understand. In summary, meta-analysis is a method of testing whether findings from multiple studies involving bivariate analysis are homogeneous or heterogeneous, or in other words, do they agree or disagree in terms of the direction of association and effect size? If the findings are homogeneous, proponents of meta-analysis then argue that it is possible to calculate a truer estimate of the effect size utilizing the data from two or more studies. The meta-analyst is not averaging the findings but rather treating data from multiple studies as if they were all part of a single study. Given enough descriptive statistics in the published report, such estimates can be calculated without requiring access to the actual dataset.

This last part of the process is where opponents question the validity of the method, suggesting that data can only be properly interpreted within the context of how the observations were initially gathered (Hunter & Schmidt, 1990). However, such arguments provide means of their own refutation by defining the conditions under which meta-analysis can be considered valid. If subject populations are given the same tests or interventions using identical measures under similar conditions, then one may logically accept that multiple tests will yield a truer representation of a bivariate relationship, just as drawing multiple samples of cards with numbers on them from a hat will yield a truer estimate of the mean of all the numbers in the hat. Therefore, the selection of variables and effect size estimates to be considered when planning to conduct meta-analysis is vital in that it will limit the number of possible studies that can be included.

Rosenthal (1991) outlines a large number of effect size estimates that can be used in meta-analysis. Unfortunately, a number of these estimates are dependent on the scale of the variables in question. Even variables originally based on the same operational definition are sometimes rescaled for the purpose of a given study. To overcome this difficulty, G. V. Peckham Glass (as cited by Hedges & Olkin, 1985) proposed using scale-free estimates of effect size. Popular scale-free estimates include Cohen's d and Glass's alpha, but these measures are specifically designed for use in experimental or comparison studies where at least two groups of subjects are involved. Many studies in LIS are descriptive in nature and do not involve the use of control groups.

Effect size estimates that are not scale-free (for example, correlation coefficients) are susceptible to bias. Small sample sizes will cause wide variability in estimates across studies. Also, range restriction of indicators for the dependent or independent variable may reduce the value of the estimate. For example, a correlation coefficient based on a measure using a seven-point Likert scale is likely to be lower than that obtained from a measure using a four-point scale. The best way to avoid criticism when using such

estimates is to only compare variables across studies that have been measured using the same scale.

Such practice may severely limit the number of studies one may include in meta-analysis. For example, Saxton (1997) encountered a number of problems when looking for repeated measures in evaluation studies of reference service performance.

> Out of fifty-nine studies, forty-two use reference accuracy as an outcome variable, but of those only twenty measure accuracy on the same scale. . . . Out of those twenty studies, only five reported the correlation coefficients between reference accuracy and a multitude of independent variables . . . [Of these], three studies sample fewer than twenty subjects. (p. 274)

The situation did not improve when examining independent variables. Saxton goes on to explain that he identified 38 concepts operationalized in the form of 162 different measures. Of those 162 variables, only 10 were repeated in more than one study.

Alternatively, the amount of error resulting from comparing variables of different scales may be small, and each future meta-analyst will have to assess the extent of the possible threat to validity. When introducing a method relatively new to the discipline, future researchers are encouraged to adopt a conservative approach until acceptance is more broadly attained.

Saxton (1997) articulated that the process for comparing and recalculating effect size estimates across studies requires three steps. First, the researcher must test the homogeneity (similarity) of significance levels across studies. If the significance levels for the findings in each respective study are not homogeneous, then the findings from each sample are contradictory. It is then inappropriate to combine the findings since they are not indicating consistent conclusions. Next, the researcher must test for the homogeneity of effect size estimates across studies to determine if it is appropriate to derive a new estimate from them. For example, if for a given pair of variables one study indicates a strong association and another study indicates a weak association, the researcher cannot simply "split the difference" and declare that the combined findings indicate a moderate association. Neither study suggested that the association was moderate; the samples exhibited conflicting characteristics (Hedges & Olkin, 1985). Finally, once homogeneity has been established, the researcher can calculate a new effect size estimate and associated significance value. Studies that employ larger sample sizes are weighted so as to give them greater emphasis in the actual calculations (Matt & Cook, 1994).

Meta-analytic techniques are controversial because they are susceptible to numerous threats to validity. First, publication bias, as discussed earlier, is one danger encountered by the researcher. Frequently, studies that do not yield significant findings are not reported. Second, range restriction limits the ability to compare results across studies. Third, failure on the

part of investigators to note the number of missing cases for each variable contributes to error in meta-analysis since both significance levels and effect size estimates are strongly influenced by the number of subjects being examined. Fourth, lack of reliability in measurement and coding always threatens to invalidate the conclusions for all analyses. Researchers performing meta-analyses must apply strict quality control by excluding any studies that fail to meet methodological standards or appear to sample imprecisely (Matt & Cook, 1994).

Many different sources provide a wealth of technical detail on how to design a meta-analysis and perform the necessary calculations. Within LIS literature, Ankem (2005) offers perhaps the most sophisticated discussion of meta-analysis. She provides an overview of the three dominant methodological approaches to meta-analysis: the Hedges and Olkin approach that employs scale-free estimates of effect size estimates, the Rosenthal and Rubin approach that recommends transformation of effect size estimates to standard scores, and the Hunter and colleagues approach that attempts to correct for various sources of error in individual studies. This is followed by an illustrative example of a meta-analytic study of factors affecting information needs of cancer patients. An earlier study by Saxton (1997) provides a narrower, simpler example utilizing the Rosenthal and Rubin approach in a meta-analysis of studies of reference service quality. Both Ankem and Saxton cite Rosenthal's (1991) handbook, *Meta-analytic Procedures for Social Research,* as a useful and relatively accessible technical source for providing guidance on which calculations to use and addressing methodological concerns.

Literature Review

A search in *Library and Information Science Abstracts* (*LISA*) reveals that not only is the methodology rarely applied, but that the term itself, meta-analysis, rarely appears. Conducting a search for the terms *meta-analysis* or *metaanalysis* in any field yielded references to only 51 journal articles, and a search for the phrase *research synthesis* yielded only 1 article. Of these 52 articles, only 21 appear in LIS-oriented journals, while the other references are meta-analytic studies in the disciplines of communication, education, or human-computer interaction. While these studies all involve information and technology and may be of interest to LIS researchers, this review will focus on studies that appear in the LIS literature.

Meta-analysis has a long history in medicine, and health science librarians are perhaps the LIS professionals most familiar with the technique. Schell and Rathe (1992) have the earliest, though also brief, mention of the term *meta-analysis* in *LISA* when describing the method as a "quantitative procedure for combining results of clinical trials" (p. 219); they further note the important role that librarians will play in helping researchers conduct extensive literature reviews as this method gains in popularity. Over the

past ten years, this theme has been echoed by many others discussing the challenges for medical researchers faced with large retrieval sets, the difficulties encountered in conducting exhaustive searches for the purpose of meta-analysis, and the ability of librarians to assist researchers (McKibbon & Dilks, 1993; Smith, Smith, Stullenbarger, & Foote, 1994; Mead & Richards, 1995; Smith, 1996; Timpka, Westergren, Hallberg, & Forsum, 1997; Johnson, McKinin, Sievert, & Reid, 1997; Yamazaki, 1998; Royle & Waugh, 2004; Demiris et al, 2004).

Interest in the method as a means to investigating research problems in LIS began to grow in the early 1990s. Trahan (1993) discussed the feasibility of meta-analysis in LIS and attempted to inform researchers about the potential of this methodology. Harsanyi (1993) suggested that studies of collaborative authorship would be a good topic for meta-analysis because of the complex relationship between collaboration and productivity.

The first published meta-analysis performed by an LIS researcher appeared in 1996. Salang (1996) used Glass's techniques in studying the relationship between user needs and options for retrieving information. However, the study was not published in a widely read journal and is not frequently cited.

The following year, Saxton (1997) performed a meta-analysis of reference service evaluation studies. The primary research question was to determine what factors predicted levels of accuracy in answering questions. Out of fifty-nine studies taking place over a thirty-year period from 1965 to 1995, only seven were eligible for inclusion in the meta-analysis because they reported sufficient descriptive data and used the same measures. Findings indicated that factors such as collection growth, library budget, and hours of operation consistently exhibited a positive moderate association with response accuracy. However, the greater value of this study was to provide a step-by-step demonstration of how to conduct a meta-analysis and discussion of methodological concerns such as publication bias, quality standards, requisite sample size of studies, the need for replication of previous studies, and the need for greater uniformity in reporting research.

To model the desirable practice he was advocating, Saxton (1997) provided sufficient statistical data to enable later researchers to include his work in future analysis. This action was clearly validated four years after publication when a doctoral student, Rafael Merens, at the University of Havana, Cuba, re-analyzed Saxton's work for his dissertation. Merens examined the same seven studies using a different meta-analytic approach to optimize the value of studies with small samples, resulting in alternative estimates of combined effect size (Merens & Morales, 2004).

Hwang and Lin (1999) reported the results of a meta-analysis examining the effect of information load (defined in terms of both information diversity and repetitiveness) on decision quality of managers as reported in bankruptcy prediction experiments. The meta-analysis compared findings

from thirty-one experiments reported in eighteen studies but excluded several studies "that did not report requisite data" (p. 215). In conclusion, the researchers noted the success of meta-analysis in clarifying inconsistencies in the research record: "This meta-analysis has found clear evidence of the detrimental effect of information load on decision quality. Results showed that decision quality suffers with an increase in either the diversity or repetitiveness of an information cue set. The findings help to reconcile the inconsistent evidence reported in the bankruptcy prediction literature" (p. 216). Their article ends with a discussion of the implications for both information suppliers and information retrieval.

Wantland et al. (2004) published a complex, large-scale meta-analysis concerning how the medium of an intervention (Web-based vs. non-Web-based) influences the behavior change of an individual with a chronic condition. This study may be the first attempt in the medical library literature to apply meta-analysis to an information research problem rather than a clinical research problem. In preparation, the research team conducted an extensive systematic review of the literature (see McKibbon's article in this issue for more information on systematic reviews). Each study was rigorously reviewed for its suitability for inclusion in the meta-analysis.

> The compliance to standards for the studies is based on five criteria: (1) study design; (2) selection and specification of the study sample; (3) specification of the illness/condition; (4) reproducibility of the study; and (5) outcomes specification and the measurement instruments used/validity and reliability of documentation of instruments. The sum of the variables result in a total score ranging from 0 to 18 . . . Only studies with a quality documentation score of 12 or greater were retained for the meta-analysis. (Wantland et al., 2004, p. 3)

The study used a scale-free estimate of effect size, Hedges *d,* to assess the impact of intervention medium on user behavior. The findings conclusively demonstrated that Web-based interventions were consistently more effective than other interventions, although the actual effect size varied widely and was not homogeneous across studies.

Ankem (2005) presents a more thorough, detailed discussion of methodology in her meta-analysis of factors affecting information needs among patients. After discussing the merits of three different statistical approaches to meta-analysis, she notes that the procedure is rarely used in LIS: "The reasons for the lack of use of meta-analysis in LIS may be attributed to the difficulty in accumulating results involving variables related to the same research problem across studies and the lack of appropriately measured variables related to the same research problem across studies so that the results can be combined meaningfully" (p. 165). The results of the meta-analysis based on four studies indicated that the age of individuals has a negative association with their need for information, possibly suggesting that older individuals are more susceptible to information overload, or

may intentionally avoid seeking information about their medical condition, than younger individuals. One particular strength of her study is the use of studies conducted in fields other than LIS to investigate questions about information behavior. This example suggests that meta-analysis may be a useful vehicle to expand disciplinary knowledge in LIS by building on the research enterprise of "larger" fields (those with more researchers and more grant funding).

On occasion, researchers have used the term *meta-analysis* when only referring to the idea of aggregating findings across studies rather than actually performing the statistical analyses conventionally associated with the term. Haug (1997) reported on a study that utilized what he described as a meta-analytic procedure. The purpose of the study was to examine physician's preferences for using different types of information sources to answer questions in their clinical practice. Unfortunately, he encountered the same difficulties in finding suitable studies to consider.

> Comparative analysis of the twelve selected studies was limited by their dissimilar research questions, research instruments, and reportorial formats . . . Unfortunately, the published findings of the research described in this paper do not permit rigorous statistical meta-analysis. Conventional meta-analysis marshals evidence for or against relations among variables common to several studies by combining results of significance tests or statistics which measure strength of relationship. The twelve investigations analyzed in this study neither share a common hypothesis nor test for relations among a common set of variables. (p. 225)

Haug settled for aggregating data on ranking physicians' preferences since he did not find any study that tested bivariate relationships. While Haug was conscientious in his use of the term, others have been less concerned. Olson and Schlegl (2001) describe their investigation of critiques of subject access standards in the classification literature as a "meta-analysis" although the only quantitative evidence they present are percentages of topics appearing in ninety-three articles.

Despite these individual efforts, meta-analysis has largely been underutilized in LIS. Hjorland (2001) wrote a letter to the *Journal of the American Society for Information Science and Technology* lamenting that meta-analysis was being neglected by information scientists and arguing that meta-analysis was a valuable research method and also "an expansion of the professions [*sic*] possibility in relation to what should be our core competence: document searching/information retrieval" (p. 1193). However, as has been demonstrated repeatedly in the above review, issues of consistency, replication, and adequate reporting must also be resolved before meta-analysis can be more widely applied.

Recommendations for Reporting Research

The ability to conduct a meta-analysis is dependent upon the consistency with which earlier studies report findings. As discussed at the beginning of this article, it is ironic that stringent rules exist for governing the style of citations and a complex code administers the creation of bibliographic records, yet no commonly recognized standards exist for reporting the results of research in LIS. Saxton (1997) proposed a set of five minimum standards for reporting quantitative research studies that use Pearson's correlation coefficient for bivariate analysis. In response to Ankem's (2005) criticism of this narrow approach to meta-analysis, these standards are amended here as follows to accommodate a broader range of statistics:

1. Include the operational definition of every variable mentioned in the article. In some cases, such as survey research, the simplest way to do this may be to include a copy of the instrument (to save space in the journal, some items such as demographic questions may be omitted, and the instrument may be reformatted).
2. For every variable mentioned in the article, list the mean, minimum, maximum, and standard deviation. This data can be easily summarized in a short table in an appendix to the article.
3. List the number of responses for each variable. If the variable has missing cases, list the total number of subjects available for that variable. This data could also be included in the aforementioned table.
4. When describing bivariate relationships, include the precise level of significance (for example, p) associated with a given statistic for effect size (for example, Pearson's r) rather than just truncating (for example, $p < .05$). This enables the meta-analyst to calculate more accurately a significance level associated with the newly derived effect size based on multiple studies. Significance is an arbitrary level based on the degree of confidence the researcher is seeking in a given study and may often vary for studies using the same measures and methods.
5. When bivariate relationships are found to be insignificant, list the precise value of p rather than simply noting that the results were not significant. Significance is closely related to sample size, and meta-analysis utilizes larger samples by interpreting findings from multiple studies.
6. Explicitly describe the population and the unit of analysis for each variable within the population (for example, in a study of reference service, Saxton [2002] gathered observations at the library, librarian, and service transaction level). Findings across studies cannot be compared if they use different units of analysis. To apply group-level observations to individuals is known as the ecological fallacy, and to apply individual-level observations to groups is the reductionist fallacy (Schutt, 2004). Such errors result in intraclass correlation, an error that masks the

true effect size between two variables by confounding group-individual relationships.

Of course, the primary objective for researchers is to explain the phenomena they are observing and what it means in terms of expanding disciplinary knowledge and improving teaching and practice. Few researchers set out with the goal of making meta-analysis easier to perform. However, scientific research is a cumulative process where advances are made through multiple investigations over time. Investigators who follow the above guidelines will encourage that process and potentially increase the impact of their own work as exemplified in the relationship between Saxton (1997) and Merens and Morales (2004).

While consistent reporting is the first issue to overcome, the second problem is the lack of consistency in measuring concepts over time. Investigators have not been using the same operational definitions either through oversight (lack of awareness of previous studies) or intention (a belief that previous studies used poor measures). Until some consensus is reached on what definitions and indicators are best to use for the significant concepts in given problem areas, repetition of tests across multiple studies will rarely occur. In terms of quantitative research, this will retard the maturation of the discipline by preventing the accumulation of large datasets and enabling new researchers to build upon the foundation laid by experienced researchers. This may also discourage new researchers from pursuing quantitative methods as a possible means of investigation for the questions that interest them.

As a final thought, the Internet has provided a platform to make it easier to perform meta-analysis than at any other time as scholars no longer view the refereed journal article as the sole means for disseminating information about their research. As journal editors review papers with an eye to cutting out "extraneous" material to conserve pages, the World Wide Web makes it possible to share tables of variables, statistics, copies of instruments, and any other information that would be of use to colleagues investigating the same research questions. In some cases, individual researchers may now provide their actual dataset to others (subject to regulations governing the privacy concerns of human subjects). However, scholars also have many good reasons to restrict the nature of access to their data, primarily to retain control of how the data is used and how findings are interpreted and presented. Likewise, releasing instruments to the public before conducting any reliability testing or cross-validation of the different variables may only result in the repeated use of poor measures. Reporting research findings according to the recommendations given above provides a "middle road" between providing total access to data or instruments and controlled sharing that enables researchers to receive peer feedback, facilitate meta-analysis, promote research synthesis, and still maintain ownership and control of their creative work.

References

Ankem, K. (2005). Approaches to meta-analysis: A guide for LIS researchers. *Library & Information Science Research, 27*(2), 164–176.

ARL Committee on Statistics (1997). *Interim guidelines for counting materials stored in library storage centers.* Retrieved October 15, 2005, from http://www.arl.org/stats/arlstat/storage.html.

Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate.* Princeton, NJ: Council of Learned Societies.

Demiris, G., Folk, L. C., Mitchell, J. A., Moxley, D. E., Patrick, T. B., & Tao, D. (2004). Evidence-based retrieval in evidence-based medicine. *Journal of the Medical Library Association, 92*(2), 196–199.

Harsanyi, M. A. (1993). Multiple authors, multiple problems—bibliometrics and the study of scholarly collaboration: A literature review. *Library & Information Science Research, 15*(4), 325–354.

Haug, J. D. (1997). Physicians' preferences for information sources: A meta-analytic study. *Bulletin of the Medical Library Association, 85*(3), 223–232.

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Hjorland, B. (2001). Why is meta analysis neglected by information scientists? [letter to the editor]. *Journal of the American Society for Information Science & Technology, 52*(13), 1193–1194.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage Publications.

Hwang, M. I., & Lin, J. W. (1999). Information dimension, information overload and decision quality. *Journal of Information Science, 25*(3), 213–218.

Johnson, E. D., McKinin, E. J., Sievert, M., & Reid, J. C. (1997). An analysis of objective year book citations: Implications for MEDLINE searchers. *Bulletin of the Medical Library Association, 85*(4), 378–384.

Matt, G. E., & Cook, T. D. (1994). Threats to the validity of research synthesis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 503–520). New York: Russell Sage Foundation.

McKibbon, K. A., & Dilks, C. W. (1993). Panning for applied clinical research gold. *Online, 17*(4), 105–108.

Mead, T. L., & Richards, D. T. (1995). Librarian participation in meta-analysis projects. *Bulletin of the Library Medical Association, 83*(4), 461–464.

Merens, R. A., & Morales, M. M. (2004). *Los metanalisis: Aproximaciones utiles para su comprension.* La Colaboracion Cochrane en Cuba, Parte VII. Retrieved November 1, 2005, from http://eprints.rclis.org/archive/00002680.

Olson, H. A., & Schlegl, R. (2001). Standardization, objectivity, and user focus: A meta-analysis of subject access critiques. *Cataloging & Classification Quarterly, 32*(2), 61–80.

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* Beverly Hills, CA: Sage Publications.

Royle, P., & Waugh, N. (2004). Should systematic reviews include searches for published errata? *Health Information & Libraries Journal, 21*(1), 14–20.

Salang, M. M. C. (1996). A meta-analysis of studies on user information needs and their relationship to information retrieval. *Journal of Philippine Librarianship, 18*(2), 36–56.

Saxton, M. L. (1997). Reference service evaluation and meta-analysis: Findings and methodological issues. *Library Quarterly, 67*(3), 267–289.

Saxton, M. L. (2002). *Understanding reference transactions: Transforming an art into a science.* San Diego, CA: Academic Press.

Schell, C. L., & Rathe, R. J. (1992). Meta-analysis: A tool for medical and scientific discoveries. *Bulletin of the Medical Library Association, 80*(3), 219–222.

Schutt, R. K. (2004). *Investigating the social world: The process and practice of research.* Thousand Oaks, CA: Pine Forge Press.

Smith, J. T. (1996). Meta-analysis: The librarian as a member of an interdisciplinary team. *Library Trends, 45*(2), 265–279.

Smith, J. T., Smith, M. C., Stullenbarger, E., & Foote, A. (1994). Integrative review and meta-analysis: An application. *Medical Reference Services Quarterly, 13*(1), 57–72.

Timpka, T., Westergren, V., Hallberg, N., & Forsum, U. (1997). Study of situation-dependent clinical cognition: A meta-analysis and preliminary method. *Methods of Information in Medicine, 36*(1), 44–50.

Trahan, E. (1993). Applying meta-analysis to library and information science research. *Library Quarterly, 63*(1), 73–91.

Wantland, D. J., Portilla, C. J., Holzemer, W. L., Slaughter, R., & McGhee, E. M. (2004). The effectiveness of Web-based vs. non-Web-based interventions: A meta-analysis of behavioral change outcomes. *Journal of Internet Medical Research, 6(*4). Retrieved May 30, 2006, from http://www.jmir.org/2004/4/e40/.

Yamazaki, S. (1998). MEDLINE searching for evidence-based medicine. *Journal of the Japan Medical Library Association, 45*(4), 402–405.

Matthew L. Saxton is an assistant professor at the Information School of the University of Washington. His primary research interests are question-answering behavior, intermediation, and the evaluation of information services. His book *Understanding Reference Transactions* explains the use of hierarchical linear modeling to investigate the factors that contribute to success in responding to reference questions in public libraries. He received the 1997 Methodology Award from the Association for Library and Information Science Education for his application of meta-analysis to evaluation studies of reference service. He earned his Ph.D. from the University of California, Los Angeles.

Observation: A Complex Research Method

Lynda M. Baker

Abstract

As an ethnographic research method, observation has a long history. The value of observation is that it permits researchers to study people in their native environment in order to understand "things" from their perspective. Observation requires the researcher to spend considerable time in the field with the possibility of adopting various roles in order to gain a more comprehensive understanding of the people being studied. A variety of techniques are used to collect data. Gaining access to the group and leaving the field are two important factors that need consideration. Other areas of concern involve ethical problems, as well as validity and reliability issues. Until recently, few library and information science (LIS) studies have included this method; however, observation is gaining favor as LIS researchers seek to understand better the role of information in people's everyday lives.

Introduction

As an ethnographic research method, observation seems to have no specific beginning. While some researchers found indications of its use in ancient times, others have pointed to the late nineteenth and early twentieth centuries, when anthropologists starting "collecting data firsthand" (Atkinson & Hammersley, 1994, p. 249). Describing it as the "bedrock source of human knowledge" about the "social and natural world," Adler and Adler (1994) stated that Aristotle used observational techniques in his botanical studies on the island of Lesbos and that Auguste Comte, the father of sociology, listed observation as one of the "four core research methods" (p. 377).

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 171–189
© 2006 The Board of Trustees, University of Illinois

In the current research environment, its status seems to have changed, leading Adler and Adler to question whether observation is a research method "in its own right" or "a stepchild to its more widely recognized offshoot: participant observation" (1994, p. 378). Further confusing the picture is the variety of labels (for example, observation, participant observation, or ethnography) that seem to be used interchangeably by researchers to describe what was once called simply "observation." Finally, in some research methods textbooks and articles, observation has been described as a research method as well as a data collection method (Powell & Connaway, 2004; Williamson, 2000; Pearsall, 1970). Williamson prefers to categorize observation as a data collection technique because it can be used in a variety of research methods.

Observation is a complex research method because it often requires the researcher to play a number of roles and to use a number of techniques, including her/his five senses, to collect data. In addition, despite the level of involvement with the study group, the researcher must always remember her/his primary role as a researcher and remain detached enough to collect and analyze data relevant to the problem under investigation. The purpose of this article is to describe in some depth the types of roles a researcher can assume during an observational study. In addition, an overview of some of the characteristics unique to observational research, as well as validity and reliability and ethical issues, are addressed. Interspersed throughout the article are some examples of LIS studies in which the observation method has been used. Two topics are not covered in this article. The first topic is structured observation, which Glazier defined as a "qualitative research method" in which "pre-determined categories are used to guide" (1985, p. 105) the recording of activities undertaken by people in their natural environments. Because the role of the observer is limited to recording events, it is outside the scope of this article. Analysis of qualitative data has been covered in detail in a number of books (see, for example, Strauss & Corbin, 1990; Spradley, 1980) and therefore will not be covered in this article.

At this point, it is also important to mention the difficulty one encounters searching for studies that have used this method in either *Library Literature* or *Library and Information Science Abstracts* (*LISA*). Some researchers do not specify what role they played. For example, in her study of janitors, Chatman (1990) does not indicate the role she adopted. This practice leads to either broad subject headings or to the complete absence of indexing terms applied to observational studies. As part of their study of research method trends in the literature on human information behavior (HIB), McKechnie, Baker, Greenwood, and Julien (2002) examined how *Library Literature* and *LISA* indexed the methods used in 247 HIB articles published from 1993 to 2000 in seven international, peer-reviewed journals. Of the 247 articles, 152 articles were found in *Library Literature* and, of these, only "39 (26%) were indexed by at least one method term" (p. 123). LISA had indexed 178

articles, but even fewer (32 or 19 percent) articles "were indexed by at least one method term" (p. 123). Furthermore, both indexes were found to use terms that are too broad to be helpful to researchers who are searching for articles in which a particular method has been used. These results reveal the challenge of retrieving studies on specific methods.

Definition of Observation

Definitions of *observation* per se are difficult to find in the literature. Gorman and Clayton define observation studies as those that "involve the systematic recording of observable phenomena or behaviour in a natural setting" (2005, p. 40). Other authors define observation within the broader context of ethnography or the narrower one of participation observation. What is consistent in the definitions, however, is the need to study and understand people within their natural environment. Spradley wrote that participation observation "leads to an ethnographic description" (1980, p. vi). He defined ethnography as the "work of describing a culture" with the central aim of understanding "another way of life from the native point of view" (p. 3). Chatman defined ethnography as a method that allows the researcher to get an insider's view through observation and participation "in social settings that reveal reality as lived by members of those settings" (1992, p. 3). Becker and Geer defined participant observation as either a covert or overt activity "in which the observer participates in the daily life of the people under study . . . observing things that happen, listening to what is said, and questioning people, over some length of time" (1970, p. 133). To observe people in their natural settings, there are a variety of roles researchers can adopt. The roles and how they have changed over time are described below. Where possible, examples of LIS studies are included.

Roles of the Researcher

Roles have been defined as "the characteristic posture[s] researchers assume in their relationship" with the people whom they are studying (hereafter referred to as "insiders") (Chatman, 1984, p. 429). In his article on roles in field observations, Gold (1958) credited, and expanded on, Buford Junker's typology of four roles researchers can play in their efforts to study and develop relationships with insiders, including complete observer, observer-as-participant, participant-as-observer, and complete participant (p. 217). More recently others, such as Spradley (1980) and Adler and Adler (1994), have proposed slightly different roles or used different terms than did Gold, as will be discussed below.

While Gorman and Clayton described Gold's four roles as "a range of flexible positions in a continuum of participatory involvement" (2005, p. 106), not everyone has to start as a complete observer. The adopted role depends on the problem to be studied, on the insiders' willingness to be studied, and on the researcher's prior knowledge of or involvement in the

insiders' world. Going into a new environment may require the researcher to adopt the role of complete observer, whereas studying a group in which she/he is already a member allows the researcher to adopt the complete participant role. What is important is that the researcher assumes an appropriate, fluid role—one that allows her/him to observe intimately the everyday life of the insiders (Chatman, 1984; Carey, McKechnie, & McKenzie, 2001).

Nonparticipation

This role, described by Spradley (1980), involves no level of involvement with insiders. The researcher is not present on the scene but rather can "observe" from an entirely different environment. Transaction log analysis (TLA) is an example of this type of observation. In his article Davis described TLA as a "non-intrusive method for collecting data from a large number of individuals for the purpose of understanding online-user behavior" (2004, p. 327). Using TLA he focused on the American Chemical Society's servers to determine how chemists at Cornell University located information. Moukdad and Large analyzed over 2,000 search strategies submitted by users to WebCrawler to determine query characteristics and also to try "to understand how these users view the Web" (2001, p. 350). In her study, Thompson (2003) used a screen viewer to watch, from another room, the interaction of college students as they tested the library's new Web site. While this role has advantages and is effective for some LIS studies, it does not allow for any in-depth understanding of people's behavior in their own world.

Complete Observer

Gold's (1958) complete observer and Gorman and Clayton's (2005) unobtrusive observer play the same "passive" role as described by Spradley (1980). In this role, the researcher is present on the scene but, according to these three authors, does not participate or interact with insiders to any great extent. Her/his only role is to listen and observe. Within this role, lesser ones are adopted to allow the researcher to be invisible while, at the same time, ubiquitous in order to eavesdrop (Pearsall, 1970). One advantage of this role is that the researcher can remain completely detached from the group. Detachment, however, is also a major disadvantage because it could prevent the researcher from hearing entire conversations or grasping the full significance of an information exchange. She/he cannot ask insiders any questions to "qualify what they have said, or to answer other questions his observations of them have brought to mind" (Gold, 1958, p. 222). In addition to eavesdropping, a complete observer can collect data through videotaping, audio-taping, or photographing insiders (Adler & Adler, 1994), all of which have ethical implications. Given its limitations, Gold (1958) stated that complete observer is more often used as a subordinate role to other dominant ones. He conceded, however, that this role

may be an important starting point for future observations and interactions when the researcher assumes other roles.

Although this role may not seem ideal in one's quest to understand insiders, it has its value and is often used in conjunction with other data collection techniques. A few LIS examples are provided to illustrate its usefulness. Given and Leckie used an "unobtrusive patron-observation survey, called 'seating sweeps'" in their study of people's use of public library space (2003, p. 373). They developed a "seating sweeps checklist" (p. 375) and walked through the library three times a day at different intervals to observe how people were using various spaces. Using unobtrusive participant observation, as well as audiotapes of their verbal comments and exchanges, McKechnie (2000) observed the behavior of four-year-old girls in a public library. In addition, she collected a written diary from each girl's mother. Radford (1998) studied college students' decisions to approach reference librarians. For thirty-seven hours she unobtrusively observed students and recorded the nonverbal behaviors of both librarians and clients on a structured data collection form. She also interviewed the students. In his study of people with an autoimmune disease, Carey (2003) observed members of a support group during their meetings, listened to them, and observed their interactions. He also interviewed twenty-five members of the group.

The next example demonstrates that the role of complete observer may be the only permitted way to conduct a study. The author (Baker) and her colleague (Case) wanted to interview street-level female sex workers to ascertain their health concerns (Baker, Case, & Policicchio, 2003). They were restricted, however, to the role of complete observer by outside forces, namely the human investigation committee of their university and the outreach agency with whom they were working. The former required the researchers to obtain signed informed consent from the participants, while the latter felt this procedure would inhibit the agency's work with the women. Thus, the researchers had to gather information about health issues by listening to the conversations between the sex workers and the volunteers who distributed supplies to them.

Observer-as-Participant

This role, as described by Gold (1958) and Pearsall (1970), includes more observation than participation. The researcher who adopts this role advances very slightly in her/his involvement with the insiders. While still mostly involved in observing, she/he may conduct short interviews. Unlike the covert activity that is typical of the complete observer, in this role the researcher's identity can become more overt as it becomes known to more of the insiders. The researcher, however, should remain "strongly research oriented" and "not cross into the friendship domain" (Adler & Adler, 1994, p. 380).

Pearsall (1970) described two advantages to this role. First, insiders may be more willing to talk to "attentive strangers" than they would be to talk to

people with whom they are more familiar. Second, there is less "temptation either for the observer to go native or for the natives to try to include him permanently in their lives" (p. 342). The downside of this role is that the brief encounters with insiders limit "opportunities for gaining knowledge of total situations" (p. 342). Gold saw this role as a source of frustration to the researcher who "cannot take time to master" the insiders' "universes of discourse" (1958, p. 221). In other words, the brief interviews can contribute to misunderstandings or misconceptions of which the researcher may not be aware until it is too late to correct or address them.

Few LIS studies were found in which this role was adopted. In their study of the health problems of female street-level prostitutes described above, Baker and Case accompanied volunteers of a street outreach program. Because they were unable to speak directly to the women, they relied on the volunteers to obtain health-related information from some of the women. The observations of the researchers and the volunteers, as well as the discussions between them, provided good information about the health concerns of women who worked the streets. Carey's (2003) study of the support group (mentioned above) included his participation as a librarian before and after the meetings. In this role, he was able to observe and participate to some degree by talking to the members about their selection of library materials.

Moderate or Peripheral Membership

In 1994 Adler and Adler wrote that the roles of complete observer and observer-as- participant were no longer as popular with qualitative researchers as they had been during the mid-twentieth century (p. 380). Instead, researchers preferred "greater involvement," which included what they called "membership roles" (p. 379). Thus, new role labels appeared in the literature. Adler and Adler's "peripheral membership" seems to equate to Spradley's (1980) moderate role.

In this role the researcher wants to "maintain a balance between being an insider and an outsider, between participation and observation" (Spradley, 1980, p. 60). To accomplish this, the researcher interacts with the insiders and engages in similar activities but, according to Adler and Adler she/he does not participate in those activities "that stand at the core of group membership and identification" (1987, p. 36). They postulated two reasons for adopting this role. First, the researcher may limit involvement in the group, fearing that it will affect her/his ability to interpret the data from a detached perspective. Second, the researcher may "intentionally restrict" the level of involvement because she/he does not want to participate in the specific activities of the insiders being studied (p. 36). In their study of drug dealers, this is the role Adler and Adler assumed.

From the description of her study of older women living in Garden Towers, Chatman's (1992) role was that of peripheral membership. As Gorman and Clayton (2005) pointed out, she sat with the women on a regular basis, played cards, and ate with them. Throughout her study, however, she maintained the balance between observation and participation by not becoming involved in their daily care, that is, she did not become a member of the staff in the home. This role is similar to the one Carey (2003) played as librarian at the support group meetings. His not having the disease precluded his complete membership in the group.

Participant-as-Observer, Active Participation, Active Membership

The role that Gold (1958) called participant-as-observer, Spradley (1980) and Adler and Adler (1987, 1994) labeled "active participation" and "active membership," respectively. It is in this role that the researcher becomes more involved with the insiders' central activities but still does not fully commit to "members' values and goals" (Adler & Adler, 1994, p. 380). During this period of observation, the researcher may develop relationships with the insiders, such that they become "friends." Pearsall saw this relationship as beneficial because, as friends, the insiders can "instruct the investigator in the intricacies of their personal and social worlds" (1970, p. 343). Gold (1958), on the other hand, viewed this relationship as more problematic. First, he felt that the insider may identify too much with the researcher to continue in the role of informant and may become, instead, "too much of an observer" (p. 221). Second, the researcher may "over identify" with the insider, loose objectivity, and "go native," thus jeopardizing her/his role as a researcher/observer (p. 221).

Complete Participation

Complete participation is the ultimate level of involvement as the researcher goes native and studies a group in which she/he is already a member (Spradley, 1980; Adler & Adler, 1994). Researchers act as members, not researchers, so that they do not unnaturally "alter the flow of the interaction" (Adler & Adler, 1994, p. 380). While this role is ideal for obtaining a very good understanding of the insiders, both Gold (1958) and Spradley (1980) had reservations about researchers engaging in complete participation. In this role, the identity of the complete participant is unknown to the insiders, which can be problematic for the researcher who may become so self-conscious "about revealing his true self" that she/he becomes "handicapped when attempting to perform convincingly in the pretended role" (Gold, 1958, p. 220). Furthermore, the researcher may feel that "he has so violated his observer role that it is almost impossible to report his findings" (p. 220). Spradley agreed and cautioned that "the more you know about a situation, . . . the more difficult it is to study it as an ethnographer" (1980, p. 61).

Complete Membership

In their book on membership roles, Adler and Adler (1987) state that Gold's (1958) role of complete participant is not equivalent to their role of complete membership for several reasons. First, because the researcher and the insiders "relate to each other as status equals, dedicated to sharing in a common set of experiences, feelings, and goals" (Adler & Adler, 1987, p. 67), there is no need for the researcher to assume a covert role. Second, unlike the prohibitions in complete participation about going native, researchers adopting the complete membership role are encouraged to go native because this role enhances the data-gathering process through a sharing of information between insiders and the researcher.

In their description of complete membership, Adler and Adler state that a researcher's level of commitment varies along a continuum and that progression along this continuum "is usually associated with researchers relinquishing their involvement in and commitment to their former world and adopting the *weltanschauung,* or worldview, of members" (1987, p. 67). At one end of the continuum are researchers who, although sharing the "values, beliefs, and goals of other participants" (p. 67), do not fully join the group. At the other end are people who never return from the field.

Adler and Adler (1987) divide researchers who enter into complete membership roles into two categories: opportunistic and convert. Briefly stated, opportunistic researchers are those who are already involved in or are members of a group whom they eventually decide to study. Instead of having to bring a "pretended self" (p. 69) to the research setting, they have to "create the space and character for their research role to emerge" and examine the setting from a different perspective. In this case, the membership role precedes the researcher role. The converts, on the other hand, start as researchers whose "initial interest . . . is purely data oriented" (p. 70) but then convert to become the phenomenon. Converting may take one of two routes. Researchers may "enter the field with the express intention of making a 'good faith commitment' to becoming the phenomenon" because of their "epistemological principles, their interest in the group they are studying, or their evaluation of the pragmatic requisites for studying this group" (Adler & Alder, 1987, p. 70). Other researchers may be pressured to convert by the insiders or may be influenced by their own feelings to become a member of a group.

Problems are inherent in the complete membership role. One concerns the positive/legitimate or negative/stigma connotations of a researcher's association with the study group (Adler & Adler, 1987). Not only can researchers be contaminated by the insiders' status, they may also be stigmatized by other academics for going native. Another problem involves the consequences of the complete membership role on data gathering. Adler and Adler suggested, however, that the depth of data that can be collected in this role more than compensates for the loss of scientific detachment.

Finally, the authors acknowledged the effects the complete membership role has on the "researcher's self" (1987, p. 82). Researchers who adopt this role may find that not only have they changed, but also their relationships with others have been affected by their commitment to the insider group. Thus, the role of complete membership is not one that can be entered into lightly. No studies of LIS researchers engaging in either complete participation or complete membership were found in the literature.

Characteristics Unique to Observational Research

Observation has some aspects that are unique to this research method, including training, entering and leaving the study group, length of time in the field, sampling, and data collection techniques. Each of these topics will be described briefly below.

Training

Few general LIS research texts discuss the need for special training for those who engage in ethnographic research. Spradley (1980) states that these skills could be learned only though an apprenticeship or on-the-job training in the field. So important are these skills that he wrote two handbooks "for doing ethnography" (p. vii), including *The Ethnographic Interview* (Spradley, 1979) and *Participant Observation* (1980). In her article published in an LIS health sciences journal on the use of anthropological techniques to study the information needs of physicians, Forsythe (1998), an anthropologist, also emphasized the need for formal training:

> A word of caution: perhaps because ethnographic methods are largely qualitative in nature and are intentionally unobtrusive, people without formal training in these methods often mistakenly assume that ethnography is something that anyone can do. Doing valid and reliable ethnographic research requires considerable training and practice. (p. 407)

In their article Sandstrom and Sandstrom focus on "five misleading stances or assumptions that pervade LIS writing on qualitative research design" in the hope of clarifying "how the neglect of key issues in ethnography diminishes the value of research findings for theory building and practice" (1995, pp. 163–164). Two points in their article are relevant to the issue of training for those who wish to conduct an observational study. First, the authors attacked the naive belief that qualitative research would be better if the researchers "forgo methodological training" (p. 179). Similar to Spradley (1980) and Forsythe (1998), they state that the "proper application of qualitative methods and techniques can be achieved only by trained observers" (p. 179). Sandstrom and Sandstrom also took issue with the idea that "naturalistic inquiry . . . may begin with little or no awareness of existing literature" so that the researcher can "observe with no preconceived ideas or biases" (1995, p. 179). This view, according to the

authors, is a "flagrant violation of common sense" (p. 179). To emphasize their point, they referred to Glaser and Strauss's 1967 influential book, *The Discovery of Grounded Theory: Strategies for Qualitative Research,* wherein these authors devote one chapter to the importance of critically reading the literature. Sandstrom and Sandstrom (1995) thus suggest that researchers get a "thorough grounding in the literature" (p. 180) before they start a project because "[n]eglecting to read others' work condemns the researcher to rediscover what is already known and to repeat mistakes that could have been avoided" (p. 180).

Gaining Access and Leaving the Field

If the researcher is already a member of the group she/he is interested in studying, then gaining access is not a problem. The issues for these researchers are whether, when, and to whom to disclose oneself as a researcher (see discussion above on complete participation and complete membership; see also Labaree, 2002). Despite well-planned research and/or particular interest in a group, gaining entry is not an easy process. Time, effort, patience, and diplomacy are essential for success. In addition, maintaining that access is an ongoing process rather than a static one. A few examples from LIS studies demonstrate the difficulties researchers can encounter.

Chatman (1992) recounted her problems gaining entry in three different studies. In her research on single mothers in the Comprehensive Employment Training Act programs, she had to go through months of negotiation with city officials and site supervisors, one of whom terminated her study early for no apparent reason. When she studied janitors, Chatman had problems with the supervisors and the janitors, some of whom were "suspicious of some lady going around snooping and asking questions for some survey!" (1990, p. 5). Although gaining access to women in Garden Towers was easier, Chatman lost time starting the research project when the resident manager quit, requiring her to wait until another one was hired. She also discussed accessibility to the residents after she had gained access. The residents of Garden Towers closed their doors when they did not want to be disturbed. Although she did not violate this informal policy, Chatman noted the time lost to interviews (even prearranged ones) if a woman had closed her door.

This author (Baker, 2004) also had trouble gaining access to female vice police officers who work undercover as sex workers. It took approximately one year of negotiation with the head of the vice unit to gain access to the officers. One reason for this was that the head of the vice department changed during the negotiation period, which required starting the negotiations over with the new person. Once permission to interview the officers was obtained, there were no further problems obtaining permission to accompany and observe the officers during one of their night shifts as undercover street-level sex workers.

Observation also requires researchers to consider how to "leave the field," although, according to Labaree (2002), little attention has been paid in the literature to the process of disengagement. When the study questions have been addressed or when data saturation becomes evident, most researchers know it is time to leave. How they leave—abruptly or gradually—is the major issue they have to address. External factors, such as termination of funding, personal health, or withdrawal of permission to continue the study, may precipitate abrupt termination of the study (Jorgensen, 1989). Gradual departure may be more the norm when the researcher has adopted the complete participant or membership role. As Jorgensen pointed out, researchers may have to return periodically to get questions answered or to complete unfinished business.

Further complicating one's departure is the emotional attachment that may have formed between the researcher and the insiders and the end of relationships that have become "close and intimate . . . over lengthy periods" (Jorgensen, 1989, p. 118). In this case, Jorgensen suggested that the researcher withdraw "over a period of time so that everyone is able to prepare for the end of participant observational study" (p. 119). He also stated, however, that he has maintained contact with some of the friends he made during one of his studies. According to Adler and Adler (1987), the degree of disengagement from the study group depends on the role the researcher played. For those involved in a complete membership role, they are more likely to maintain ties with the study group than would researchers who engage in either the active or peripheral membership.

Finally, the ethical obligations to the study participants depend on the level of involvement and must be considered during the detachment period. As Labaree (2002) noted,

> Practices of strategic deceit, the tactical use of withholding information, and making conscious decisions about limiting who will read about the study's findings can follow the insider participant observer in the community long after an outsider has moved on to other research projects. These are risks that should be negotiated and carefully calculated by the insider participant observer before the study begins. (p. 115)

Length of Time in the Field

One of the unique factors of observation is the length of time in the field. Naturally, the amount of time depends on the research problem and the role assumed by the researcher. As a nonparticipant, length of time is similar to many quantitative studies. For example, in their respective transaction log analysis studies, Moukdad and Large (2001) collected data during two thirty-minute sessions in one day, while Davis (2004) collected data over a three-month period. In the other roles researchers might have to spend years in the field. Chatman, for example, spent two years studying the women in Garden Towers and two years in her study of janitors.

What is important is that the researcher have "prolonged, personal contact with events in a natural setting" (Chatman, 1984, p. 426) and play as many roles as necessary to "gain at least a comfortable degree of rapport, even intimacy, with the people, situation, and settings of research" (Jorgensen, 1989, p. 21).

Sampling

The crux of observational studies is the "who, what, where, and when" questions. Polit and Hungler (1987) divided the units of observation into two categories: molar and molecular. Molar involves observing large units of activity "as a whole," whereas the molecular approach "uses small and highly specific behaviors as the unit of observation" (p. 268). These two categories are not mutually exclusive. For example, the researcher may use the molar approach at the beginning of the study and change to the molecular one as her/his familiarity with, and understanding of, the insiders and their environment grows. Adler and Adler (1994) used the analogy of a funnel to describe this process wherein the stages of observation get progressively narrower and direct the researcher's "attention deeper into the elements of the setting that have emerged as theoretically and/or empirically essential" (p. 381).

To get rich and in-depth information, it is important for the researcher to know the best times to observe and meet with individual insiders, as well as whom she/he should interview. Extended time in the field and active participation in the group's functions increases the researcher's ability to judge these things. For example, Chatman stated that she attended many social functions at Garden Towers, including "card games and parties" (1991, p. 284). In addition, the sampling categories, such as those listed by Westbrook (as cited in Powell & Connaway, 2004; see also, Labaree, 2002), may be of some help to researchers. They include maximum variety sampling in order to make the sample as heterogeneous as possible. The researcher can also seek out insiders who "exemplify characteristics of interest" (called extreme case sampling), as well as those who have considerable experience in the group (called intensity sampling) because these people can help the researcher better understand the environment (Powell & Connaway, 2004, p. 190). Finally, the researcher may want to use snowball sampling as a way to link with others in a group. Snowball sampling is a good method to use because insiders who have been referred by a friend may be more willing to talk with the researcher. Biernacki and Waldorf (1981) identified some of the problems associated with snowball sampling that have received little attention in the literature. They dispelled the myth that snowballing is self-propelling and that once started it maintains its own momentum. Rather, the researcher "must actively and deliberately develop and control the sample's initiation, progress, and termination" (p. 143). The problems they identified include the following:

- Finding respondents and starting referral chains
- Verifying the eligibility of potential respondents
- Engaging respondents as research assistants
- Controlling the types of chains and number of cases in any chain
- Pacing and monitoring referral chains and data quality (p. 144)

Biernacki and Waldorf explained that the major problem with snowball sampling is that it is network dependent. There are two issues to consider. The first is whether the social networks formed because of the phenomenon being study and, if so, "what types of networks" have developed. Second, if the phenomenon under investigation is a "private matter," then "the problem becomes the extent to which the method will reveal the possible variations that might be extant in the population" (pp. 160–161). Thus, there is the need for the researcher to maintain "control over the referral chains" (p. 155). Other problems include "ferreting out respondents who fit the research criteria" (p. 145) and dealing with what they called "false starts," that is, the people to whom the researcher is referred turn out not to have the exact criteria for inclusion in the study (p. 149). Finally, the researcher may also need to verify participants' stories through outside sources. Although these problems can be overcome, this sampling technique requires some additional preparation and increased vigilance by the researcher to ensure that the participants meet the criteria of the study and are representative of the entire group.

Data Collection Techniques

The most common type of data collection, according to Polit and Hungler (1987), are logs and field notes. While the former are used to record daily conversations or events, field notes are "much broader, more analytic, and more interpretive" (p. 271). The researcher may choose to write, or dictate into a tape recorder, her/his field notes, which can be categorized as observational, method, theory, and personal (Chatman, 1992; Polit & Hungler, 1987). Observational notes detail what the researcher actually saw, while method notes include strategies that were "employed or that might be employed" in future observations (Chatman, 1992, p. 15). Polit and Hungler described personal notes as the researcher's "own feelings during the research process" and theoretical notes as "interpretative attempts to attach meaning to observations" (1987, pp. 272–273). Spradley (1980) called notes taken during an event the condensed version, while the expanded version is what a researcher writes after each field session. Since the key to a successful observational study is the quality of the data collected in logs and field notes (Polit & Hungler, 1987), the researcher should, according to Spradley, adhere to three principles. First, "identify the language used for each fieldnote entry" (Spradley, 1980, p. 66); in other words, identify the speaker and use "parentheses, quotation marks, or brackets" in order to have a record that "reflects the same differences

in languages usages as the actual field situation" (p. 66). The second principle is to make a verbatim record of what a person says and be able to distinguish "native terms" and "observer terms" (p. 67). Third, Spradley discussed the importance of using "concrete language" when describing observations (p. 68). Researchers should not generalize, condense, or abbreviate the details but rather "expand, fill out, enlarge, and give as much specific detail as possible" (p. 68).

In observation, the researcher uses all of her/his senses to gather information about the phenomena under study (Adler & Adler, 1994). A variety of material should also be used to enhance sensual observations. Audio-recorders can be used to tape interviews. Video-recorders or cameras can be used to record the activities of the insiders because, according to Collier and Collier (1986), cameras are an "instrumental extension of our senses" (p. 7) that may help researchers to "see more and with greater accuracy" (p. 5). In her multimethod study of hobby cooks that included "secondary research, interviews . . . and the unobtrusive analysis of sites," Hartel took 125 photographs to "capture the titles of books or file tabs with subject headings" (2003, p. 235). Other material such as minutes of meetings, memoranda, letters, magazines, or newspaper articles can also expand one's understanding of the study group. Spradley (1980) also mentioned making maps to record observations. Given and Leckie "mapped and photographed the visual space on all floors" of both libraries they studied "to document the location of furniture and equipment" in order to create the "seating sweeps checklist" (2003, p. 375).

Ethical Issues in Observation

One of the major factors associated with observational studies is ethics. While observation is generally seen as the least intrusive data collection method, it can also be an abuse of an individual's privacy (Adler & Adler, 1994; Jorgensen, 1989; Chatman, 1992). Jorgensen argued, however, that unlike scientific research, "participant observation does *not* have human subjects" (p. 28; emphasis in original) because the people with whom the researcher interacts are not subject to any experiment. While acknowledging that researchers are responsible for their actions, he stated "the researcher is not necessarily obligated to inform people of research intentions, or even protect them from possible harmful consequences" (p. 28). In today's research environment, the institution review boards (IRBs) of most institutions would not agree with his views. As Adler and Adler (1994) pointed out, universities that receive government funding have IRBs that guide research on human participants. Their policies have "outlawed disguised research" (Adler & Adler, 1994, p. 389), which may explain why the complete observer and observer-as-participant roles, as well as covert roles in complete participation, are not being used, or are frowned upon, by researchers. In addition, without sufficient justification by the researcher,

IRBs may withhold permission to photograph, videotape, or audio-tape individuals without their informed consent.

In observational research, the complexity of fieldwork in which the researcher is engaged "make[s] it difficult, if not impossible, to adopt a single set of standards," according to Spradley (1980, p. 20). He suggested researchers follow the guidelines of the American Anthropological Association, which include (1) study participants come first; (2) their rights, interests, and sensitivities should be safeguarded by the researcher; (3) participants have the right to know the aims of the researcher; (4) the privacy of the participants must be protected; (5) the participant should not be exploited or harmed in any way; and (6) reports should be made available not only to sponsors but also to the participants and the general public (Spradley, 1980, pp. 21–25).

Chatman (1992), in her book on retired women living in Garden Towers, discussed two different types of ethical dilemmas an observer can encounter. One is "*guilty knowledge,* in which the investigator is privy to confidential information, and [the other is] *dirty hands,* or a situation in which the researcher is able to correct or reveal some wrongdoing but chooses not to do so" (p. 18). Guilty knowledge, for Chatman, resulted from a confidential discussion she had with a woman who wanted to commit suicide. Chatman stated that she withheld the information from the staff and later questioned her decision: "My decision to remain silent ultimately must be attributable to my sense that her death was not harming others. She wanted the right to die and she asked that I not tell anyone. This is a haunting part of my field experience and I still wonder if I did the right thing" (p. 20). To demonstrate dirty hands, Chatman revealed why she chose not to tell the authorities about the mistreatment of a resident. First, she did not want "to risk being seen by other residents as a person who ran to the authorities, particularly since being invited to their apartment was a trusting social act" (p. 18). Her second reason related to the norms of scholarship:

> telling the authorities about that single incident did not outweigh the benefits of being silent. In other words, the first reason is related to the norms of scientific inquiry. Using this guideline, the participant observer realizes that he or she is between two different cultures: the world of persons under study and the scientific community. In order for the investigator to meet the requirements of the scientific community, a degree of objectivity in reporting data is required. (p. 18)

Validity and Reliability

As is the case with all research, researchers must address the issues of validity and reliability. In his comprehensive article on validity in qualitative research, Johnson (1997) defines validity as research that is "plausible, credible, trustworthy, and, therefore, defensible" and posits a number of strategies researchers can use to promote validity (p. 282). One threat to

validity is researcher bias that may result from selective observation, selective recording of information, or the subjective interpretation of situations. To address bias, researchers can use multiple observers, actively engage in critical self-reflection (reflexivity), or look for negative cases "that disconfirm [the researcher's] expectations and explanations" (Johnson, 1997 p. 284; Adler & Adler, 1994). In addition, Chatman used "additional methods of inquiry" (1992, p. 13), which, in her study, included an interview guide.

Johnson categorized validity as descriptive, interpretive, and theoretical and suggested strategies to promote each type. Descriptive validity "refers to the factual accuracy of the account as reported by the researchers" (1997, p. 284). He suggested "investigator triangulation" or the use of more than one investigator to collect and analyze the data (p. 283). Interpretive validity involves "accuracy in reporting the facts" or "accurately portraying the *meaning* attached by participants to what is being studied" (p. 285; emphasis in original). Strategies to improve interpretive validity include participant feedback and the use of "low inference descriptors" (that is, direct quotations) (p. 283; see also, Adler & Adler, 1994). Theoretical validity refers to "the degree that a theoretical explanation developed from a research study fits the data and, therefore, is credible and defensible" (p. 286). To promote theoretical validity, Johnson suggested that the researcher spend more time in the field. In addition, she/he can also use what Johnson called "pattern matching" (p. 283), a process that involves "predicting a series of results that form a 'pattern' and then determining the degree to which the actual results fit the predicted pattern" (p. 283). Theory triangulation would allow the researcher to examine and explain the phenomenon from different perspectives. Investigator triangulation and peer review could also help improve theoretical validity.

For Chatman (1992), validity in observational studies concerns whether the researcher is given a true picture of the phenomenon under investigation. She mentioned three types of validity: face, criterion, and construct. Face validity involves whether the observations make sense and fit into an "expected or plausible frame of reference" (p. 12). Criterion validity refers to the accuracy of findings and can be addressed by using more than one data collection technique. Chatman not only took notes but also used an interview guide (see also Adler & Adler, 1994). Finally, similar to theoretical validity is what Chatman called construct validity, which "refers to the analysis stage of field work" when the researcher determines how well the phenomenon studied fits with the conceptual framework guiding the study (p. 14).

Qualitative research is often criticized for lacking reliability. While many qualitative researchers may not be interested in generalizing their results, they must address the reality of their findings. To do so, Adler and Adler suggested that researchers should conduct their observations "systematically and repeatedly over varying conditions," that is, varying the time and the place in order to "ensure the widest range of observational consistency" (1994, p. 381).

Johnson (1997) discussed generalizability (external validity) from two perspectives. In qualitative studies, participants and the setting are not randomly selected. Furthermore, many qualitative researchers are more interested in studying "what is unique about a certain group of people, or a certain event" (p. 289). These two factors make it difficult to generalize from the sample to the population. He noted, however, that some researchers "argue that rough generalizations can be made from qualitative research" (p. 290). To do so, the group studied must be similar to the group about which one wants to generalize. Johnson (1997) suggested supplying the following information to help readers know when they can generalize:

- The number and kinds of people in the study
- How they were selected to be in the study
- Contextual information
- The nature of the researcher's relationship with the participants
- Information about any informants who provided information
- The methods of data collection used
- The data analysis techniques used (p. 290)

All this information will allow the reader not only to "make an informed decision about to whom the results may be generalized" but also to decide whether she/he would want to duplicate the study with other insiders (p. 290).

Conclusion

The literature on observation reveals how complex, challenging, and creative this research method is. Observational research differs from other methods in that it requires the researcher to have more specialized training on how to observe, what and how to record the data, how to enter the field and leave it, and how to remain detached and involved at the same time. The fact that the researcher may have to assume one or more roles is unique to observational studies. There are, however, some similarities to other research methods such as the need to plan the overall project, review the literature, and determine who will be studied and when and where (in what locations) the observations will take place. Finally, the use of one's senses, as well as other data collection techniques, make observation a more holistic type of research that allows the researcher to gain a better understanding of insiders from their own perspective. While LIS researchers are designing studies using the observation method, few have assumed the complete participant or complete membership roles. These roles might be interesting and challenging ones to assume in our efforts to understand an insider's view of the role of information in her/his everyday life.

References

Adler, P. A., & Adler, P. (1987). *Membership roles in field research: Vol. 6. Qualitative research methods.* Newbury Park, CA: Sage Publications.

Adler, P. A., & Adler, P. (1994). Observational techniques. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 377–392). Thousand Oaks, CA: Sage Publications.

Atkinson, P., & Hammersley, M. (1994). Ethnography and participation observation. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 248–261). Thousand Oaks, CA: Sage Publications.

Baker, L. M. (2004). The information needs of female police officers involved in undercover prostitution work. *Information Research, 10*(1), paper 209. Retrieved February 26, 2006, from http://InformationR.net/ir/10-1/paper209.html.

Baker, L. M., Case, P., & Policicchio, D. L. (2003). General health problems of inner-city sex workers: A pilot study. *Journal of the Medical Library Association, 91*(1), 67–71.

Becker, H. S., & Geer, B. (1970). Participant observation and interviewing: A comparison. In W. J. Filstead (Ed.), *Qualitative methodology: Firsthand involvement with the social world* (pp. 133–142). Chicago: Markham.

Biernacki, P., & Waldorf, D. (1981). Snowball sampling: Problems and techniques of chain referral sampling. *Sociological Methods & Research, 10*(2), 141–163.

Carey, R. F. (2003). *Information and narrative: A study of a self-help support group.* Unpublished doctoral dissertation. University of Western Ontario, London, Ontario, Canada.

Carey, R. F., McKechnie, L. E. F., & McKenzie, P. J. (2001). Gaining access to everyday life information seeking. *Library & Information Science Research, 23,* 319–334.

Chatman, E. A. (1984). Field research: Methodological themes. *Library & Information Science Research, 6*(4), 425–438.

Chatman, E. A. (1990). Alienation theory: Application of a conceptual framework to a study of information among janitors. *RQ, 29*(3), 355–367.

Chatman, E. A. (1991). Channels to a larger social world: Older women staying in contact with the great society. *Library & Information Science Research, 13,* 281–300.

Chatman, E. A. (1992). *The information world of retired women.* Westport, CT: Greenwood Press.

Collier, J., Jr., & Collier, M. (1986). *Visual anthropology: Photography as a research method.* Albuquerque: University of New Mexico Press.

Davis, P. M. (2004). Information-seeking behavior of chemists: A transaction log analysis of referral URLs. *Journal of the American Society for Information Science & Technology, 55*(4), 326–332.

Forsythe, D. E. (1998). Using ethnography to investigate life scientists' information needs. *Bulletin of the Medical Library Association, 86*(3), 402–409.

Given, L. M., & Leckie, G. J. (2003). "Sweeping" the library: Mapping the social activity space of the public library. *Library & Information Science Research, 25,* 365–385.

Glaser, B. G., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine.

Glazier, J. (1985). Structured observation: How it works. *College & Research Library News, 46*(3), 105–108.

Gold, R. L. (1958). Roles in sociological field observations. *Social Forces, 36*(3), 217–223.

Gorman, G. E., & Clayton, P. (2005). *Qualitative research for the information professional* (2nd ed.). London: Facet.

Hartel, J. (2003). The serious leisure frontier in library and information science: Hobby domains. *Knowledge Organization, 30*(3/4), 228–238.

Jorgensen, D. L. (1989). *Participant observation: A methodology for human studies.* Thousand Oaks, CA: Sage Publications.

Johnson, R. B. (1997). Examining the validity structure of qualitative research. *Education, 118*(2), 282–292.

Labaree, R. V. (2002). The risk of "going observationalist": Negotiating the hidden dilemmas of being an insider participant observer. *Qualitative Research, 2*(1), 97–122.

McKechnie, L. (E. F.) (2000). Ethnographic observation of preschool children. *Library & Information Science Research, 22*(1), 61–76.

McKechnie, L. (E. F.), Baker, L. M., Greenwood, M., & Julien, H. (2002). Research method trends in human information literature. *New Review of Information Behavior Research, 3,* 113–125.

Moukdad, H., & Large, A. (2001). Users' perceptions of the Web as revealed by transaction log analysis. *Online Information Review, 25*(6), 349–358.

Pearsall, M. (1970). Participant observation as role and method in behavioral research. In W. J. Filstead (Ed.), *Qualitative methodology: Firsthand involvement with the social world* (pp. 340–352). Chicago: Markham.

Polit, D. F., & Hungler, B. P. (1987). *Nursing research: Principles and methods* (3rd ed.). Philadelphia: J. B. Lippincott.

Powell, R. R., & Connaway, L. S. (2004). *Basic research methods for librarians* (4th ed.) Westport, CT: Libraries Unlimited.

Radford, M. L. (1998). Approach or avoidance? The role of nonverbal communication in the academic library user's decision to initiate a reference encounter. *Library Trends, 46*(4), 699–717.

Sandstrom, A. R., & Sandstrom, P. E. (1995). The use and misuse of anthropological methods in library and information science research. *Library Quarterly, 65*(2), 161–199.

Spradley, J. P. (1979). *The ethnographic interview.* New York: Holt, Rinehart & Winston.

Spradley, J. P. (1980). *Participant observation.* New York: Holt, Rinehart and Winston.

Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Newbury Park, CA: Sage Publications.

Thompson, S. M. (2003). Remote observation strategies for usability testing. *Information Technology & Libraries, 22*(1), 22–31.

Williamson, K. (2000). *Research methods for students and professionals: Information management and systems.* Wagga Wagga, NSW: Charles Sturt University, Center for Information Studies.

Lynda M. Baker, Ph.D., is an Associate Professor in the Library and Information Science Program at Wayne State University. She teaches courses on research methods, health sciences librarianship, and reference. Her research area is in the broad category of information needs and information-seeking behavior, with an emphasis on health-related issues. She has published articles on physician-patient communication, readability of consumer health material, and the information needs of people with multiple sclerosis. She has also co-written a book for librarians entitled *Consumer Health Information for Public Librarians.* More recently, her research focus has changed to address the health-related information needs of female, street-level sex workers.

Observations of Babies and Toddlers in Library Settings

Lynne (E. F.) McKechnie

Abstract

Participant observation, unlike the more traditional approach of querying adults about children's experiences, is identified as an appropriate and effective method for studying babies and toddlers in public library settings in order to explore these experiences from the children's own perspectives. In an observation study of eleven, thirty-minute baby storytimes conducted at two branches of a large public library system, the naturally occurring behavior of the children captured through observation field notes and audio-recording and transcription of the program successfully revealed numerous incidents of emergent literacy activities and social interaction. This article discusses the practicalities of implementing participant observation in storytime programs for very young children. Special requirements related to informed consent, the need to protect baby and toddler participants, and the challenge of gaining and maintaining access are addressed. Included is an appendix of recommended observation, child development, and research methods texts.

Introduction

Library programs for very young children (birth through two years) and their adult caregivers are common public library initiatives designed both to introduce caregivers to library resources for young children and to provide two conditions thought to enhance children's emergent literacy: a print-filled environment and "a caring adult to introduce the child to literary pleasure" (Greene, 1991, p. 7). Although much literature exists for practitioners justifying such programs and providing instructions on

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 190–201
© 2006 The Board of Trustees, University of Illinois

how to conduct them (for example, ALA, Association of Library Service to Children, 1990, 1997; Dixon & Dowd, 1993; Dowd & Dixon, 1996; Feinberg & Deere, 1995; Feinberg, Kuchner & Feldman, 1998; Flatow, 1997; Maddigan & Drennan, 2003; Nespeca, 1994), little empirical research has been conducted to delineate what actually goes on in such programs and the benefits of the programs for the children and their caregivers. It is likely that this is at least partially due to the difficulties inherent in collecting empirical data about very young children in library settings.

Infancy and toddlerhood are seen as important periods in the human lifespan. In addition to library and information science (LIS), the disciplines of developmental psychology, early childhood education, sociology, anthropology, and the health sciences have had a deep and continuing interest in very young children. While most of these disciplines recognize that babies and toddlers "take an active role in exploring the physical world and shaping their interaction with others" (Caulfield, 2001, p. 3), Greig and Taylor note that "the younger the child, the less likely the child is to be heard in research" (1999, p. 46). "Traditionally, childhood and children's lives have solely been explored through the views and understandings of their adult caretakers" (Christensen & James, 2000, p. 2). However, participant observation in naturalistic settings is emerging as a technique that is particularly well suited to studying young children in a variety of contexts. It has been identified as "particularly helpful for doing research with young children who may be unable to communicate any other way" (Greig & Taylor, 1999, p. 85). As Cohen, Stern, and Balaban note, "[c]hildren communicate with us through their eyes, the quality of their voices, their body postures, their gestures, their mannerisms, their smiles, their jumping up and down, their listlessness. They show us, by the way they do things as well as what they do, what is going on inside them" (1997, p. 6). The participant observation study described below demonstrates that this is an effective method for studying young children in library and other information settings.[1]

The Study

In order to discover what happens at library baby storytime programs and if and how these programs benefit the children who take part, we[2] conducted an exploratory participant observation study. Two sessions of baby storytime, consisting of a total of eleven, thirty-minute programs at two branches of a large public library system, were observed and audio-recorded. Interviews, both individual and focus group, were also conducted with adult participants. Data collected included observation field notes, transcripts of audio-recorded storytime sessions and interviews, and relevant documents such as program flyers and thematic booklists. Following the practices of Strauss and Corbin (1998), the data were scanned for emergent themes. Trustworthiness was ensured through strategies such as

prolonged engagement, triangulation of data collection sites, triangulation of researchers, member checking, and peer debriefing. Results of the study indicate that library storytimes provide a context in which young children are engaged in early literacy activities and social interaction and where adult participants seek, give, and exchange information (McKechnie & McKenzie, 2004). While the adult interview data was particularly useful for confirming and complementing what we learned from the children, this article focuses on the methods used to collect data directly with the babies and toddlers themselves.

Participant Observation at Storytime

Doing Storytime

A team of two researchers and one research assistant attended each program, observing before, during, and after each session. Before the program families were observed as they arrived at the library and used library facilities and services. During this time data were collected in the form of observation field notes. One researcher checked out the room where the program would be held. Careful notes were made to document the layout and set up of the room as well as any materials such as information brochures, book displays, or toys that the librarian had set out for program participants. Samples of brochures were collected and inventory lists of books and toys were made.

The actual programs themselves were audio-recorded and later transcribed. Two, and at times up to three, tape recorders were spread throughout the room to capture the program. At one of the first sessions visited in the study, one of two recorders failed, underscoring the importance of using backup equipment. More important, the use of two or three recorders was necessary as many conversations and other interactions and activities occurred simultaneously during the programs (for example, a mother speaking with a baby to direct her gaze to the librarian reading a picture book), resulting in noise levels that obscured activities in other areas of the room. During transcription, having access to two or three recordings of the same program resulted in a more complete record for that program. The research team also observed and made field notes during the program.

Anyone who has seen a baby storytime will understand just how difficult observation in this setting can be. With eight to fifteen pairs of children and caregivers (and often older siblings, friends, and relatives) and one librarian, there were a large number of participants to keep track of at any one time. Many of the children were mobile, crawling and toddling throughout the space and making it difficult to track their movement. The noise in the room made it hard to hear what was going on. To deal with these challenges, we devised a number of observation strategies. First, we always made sure to have three observers at each session. More would have been preferable

but probably too intrusive. Second, we spread ourselves throughout the room, choosing spots where we had good sight lines and covered the entire space. We divided responsibility for observation of adult-child pairs among ourselves. This allowed each observer to focus on a reasonable number of participants, usually three to five pairs. As children moved in and out of observation areas, we assumed or passed on responsibility for observing them as appropriate. Finally, while the participants we were observing were usually across the room, we also watched and listened to what was going on immediately around us. As the recorders were typically placed near an observer, this data complemented what was caught on tape. After the program was over, we again observed what families did in the library. Immediately after leaving the field, we sat down and discussed what we had experienced. This peer debriefing was important to clarify and make sense of our data and to theoretically focus our observation in subsequent sessions.

While it was impossible to observe and record everything that went on in each storytime, participant observation worked well to capture significant episodes of the children's naturally occurring behavior. The following excerpt from our field notes is typical of the many literacy events we observed.

Library 2, Session 4, Field Notes

Context: Librarian is reading a story where a double-spread illustration of an animal is followed by a double spread showing the animal making its characteristic sound.

Observation Note: Louise (8 months) is smiling in anticipation of the page turning. Librarian turns the page. Librarian and Moms roar like a lion. Louise excitedly waves her arms up and down all the while smiling broadly.

While Louise could not yet talk or even roar like the adults in the room, her smiles and body movements speak clearly to her engagement with this shared story reading. Anticipation, an important emergent literacy skill, is evident in Louise's smile. While she was not yet able to speak, Louise "roared" in her own way through her energetic arm waving. In a similar fashion, Mark and David danced their way through a story.

Library 1, Session 6, Field Notes

Observation Note: Mark (17 months) and David (15 months) are dancing around in the middle of the room. Mark is singing, twirling, and making a galloping movement. Other children move in and out of the middle of the room as the dance goes on.

Theory Note: . . . it was such a joyful thing, spontaneous and comfortable. To me it felt like magic. It's clear that both boys really enjoy their time at storytime and they express that enjoyment through their bodies.

Observation also worked well to capture the social interaction between program participants, including the youngest babies, as is evident in the following field note.

Library 2, Session 4, Field Notes

Observation Note: Thomas (8 weeks) is sitting in his Mom's lap. Daniel (6 months) leans forward and reaches out to touch his hand.

Attending a series of sessions involving the same children afforded an opportunity to observe children developing new skills. For example, during the first of six weeks a baby might watch while his mother manipulated his hands during a tickle rhyme. By the third or fourth week that child might smile and hold his hand out to be tickled when the librarian announced "let's play 'Teddy bear, teddy bear, turn around.'" The following excerpt shows how many of the children were able to learn the rhythm, conventions, and rites of storytime practices.

Library 2, Session 4, Field Notes

Observation Note: Suzanne (15 months) is carrying a nametag and she takes it over to Daniel's Mom. Suzanne looks at the nametag in her hand, then very deliberately points to the nametag Daniel's Mom is wearing. Daniel's Mom says, "You're right. It is like my nametag."

Observation proved to be an effective method for exploring what happens during storytimes from the perspective of the babies and toddlers themselves. We agree with Greig and Taylor when they say "Very young children are able to identify people, objects and places either verbally or by pointing to them" (1999, p. 78). The trick is to use a method appropriate for their developmental stage.

Getting Informed Consent

Obtaining informed consent when working with young children presents both philosophical and pragmatic problems to the researcher. Because of their legal status as minors and their limited ability to understand the research process and its potential risks, parents and guardians exercise proxy consent for them (Langston, Abbott, Lewis & Kellett, 2004; Thompson, 1992). Even when parental consent is given, researchers must take care to ensure that a child does not experience distress or any other harm. We carefully monitored the babies and toddlers whose parents allowed us to observe them at baby storytime, looking for signs of distress or discomfort. Fortunately, and possibly because the children were attending with a loving and trusted parent or caregiver, we encountered no evidence of distress on the part of any child participant. Had we done so, we would have ceased observation immediately. While some parents gave us permission to use their children's actual first names, as researchers we felt obligated to protect the privacy of the babies. In order to maintain confidentiality,

we use pseudonyms in all reports (including this one) arising from this study.

While we worked hard to collect informed consent from all child and adult participants before both storytime sessions began, the very nature of the program made this difficult. Even though pre-registration was required, this was more frequently done by telephone than in person. The requisite information sheet and consent forms were then mailed to families. Some children were registered for and attended storytime with a caregiver rather than a parent. In these cases, the researchers either contacted a parent directly or asked the caregiver to give the form to the parent and direct them to call if they had any questions. Several parents, although willing to participate in the study, forgot to bring their signed forms with them to the first session, making it necessary for them to sign new forms. Both librarians welcomed new families into the program at the last minute either just as the storytime was starting or five minutes into a program, affording us no opportunity to explain the study and get informed consent. Participating children and their caregivers often brought unexpected guests such as older siblings or visiting grandparents for whom informed consent was needed before observation could take place. We quickly learned that it was essential to have information sheets and consent forms available at all times and to designate one member of the observation team as having the responsibility of identifying new storytime attendees and seeking their consent for participation in the study. There often were one or more child/caregiver pairs present at storytime for whom we did not have signed consent forms. One grandmother, for example, did not remember to bring in the form her daughter had signed giving permission for the granddaughter to participate in the study until the penultimate session of the program. After reading a description of the fluid structure of a typical library storytime, the Research Ethics Board at our university allowed us to collect data under the proviso that we would not observe individuals for whom we did not have informed consent and that we would stop observation if anyone expressed discomfort with being present in a storytime where others were being observed. This approach worked well. Portions of the audio-recording involving such attendees were not transcribed; nor were observations and field notes made in regard to situations where the attendee interacted with other children, caregivers, or the librarian. Omission from the observation record assured that the rights of these attendees were respected.

Gaining and Maintaining Access

Carey, McKechnie, and McKenzie define access as "gaining entry to participants over a sustained time" (2001, p. 320) and describe it as "an emergent process dependent on the characteristics of the researcher, the participants, and the research context" (p. 319). Research suggests that a number of factors influence the ability of a researcher to gain access to

young children. These include gender (Pattman & Kehily, 2004; Holmes, 1998), differences in power (Robinson & Kellett, 2004), and differences in ethnicity (Holmes, 1998). In this study access was negotiated and maintained with the children in a variety of ways.

The researcher team adopted participant-observer roles. We attended the storytimes, sitting on the floor among or very near the participants. Holmes notes that with very young children "female researchers may have an inherent size advantage over male researchers because they appear smaller and perhaps less intimidating and unthreatening" (1998, p. 56). As the researchers and the research assistants were all female, gender was not a significant intervening variable during data collection. Nor did ethnicity come into play. Very few participants were members of visible minorities and all appeared to speak English as their first language, characteristics shared by the research team. In our larger study of early literacy environments, we hope to explore multicultural settings and will need to attend to and ameliorate for cross-cultural influences. The majority of caregivers in both branch libraries were the children's mothers, with the exception of one father. Both researchers are mothers and both, as professional librarians, had given a number of baby storytime programs in public libraries before becoming academics. The observation team members took care to dress like the caregiver participants. Familiarity with the setting, shared backgrounds with the adult caregivers, and making an effort to appear and act like the mothers meant that we looked very much like the other adults in the room and as such were not likely to stand apart in the eyes of the children. While there always are inherent differences in power between children and adults (Fine & Sandstrom, 1988; Graue & Walsh, 1998), by taking great care to be warm and welcoming in all our interactions with the children, we minimized this as much as possible. As can be seen through the following field note excerpt, we frequently observed incidents where the children treated us as they would any of the mothers at the program, an indication that we were successful in gaining and maintaining access.

Library 2, Session 4, Field Notes

Context: The storytime has not yet started. However, families are beginning to arrive. A basket of toys to be put out for the children at the end of storytime is on a table.

Observation Note: Suzanne (15 months) reaches into the basket of toys on the table and pulls out a sheep hand puppet. She then toddles over to Pam (a researcher), and holds out the toy toward her. Pam sings Baa Baa Black Sheep. Suzanne stays to listen, carefully looking at Pam and smiling. She toddles off at the end of the song.

Suzanne became so comfortable with Pam that she ended up "adopting" her as a caregiver for part of the final program of the storytime session.

Library 2, Session 5, Field Notes

Context: The librarian is reading Tomie DePaola's *My Halloween,* a board book.

Observation Note: Suzanne (15 months) has settled into Pam's (researcher) lap. She stays there throughout the story.

Some behaviors did differentiate us from the other adults in the room. All three observers had clipboards, paper, and pens and were almost constantly occupied writing draft field notes. None of us had a child with us. We were not the only adults with a different role —the librarian, of course, had her own distinct role and activities. The following incident provides evidence that at least some of the children were able to discern the differences between adult participants and differentiate our observation activities from the actions of the other adults.

Library 2, Session 2, Field Notes

Context: The structured part of the storytime has just ended. The librarian is moving around the room, trading cookies for the finger puppets she had passed out earlier for use with one of the rhymes. Several of the children have left their mothers and are moving around the room.

Observation Note: Samuel (14 months) wanders over to the window. His gaze is fixed on the butterflies that are hanging there. He then looks down at the recorder which is on the window ledge. He reaches up and touches the recorder. Sam then turns and looks at me. He walks over to where I am standing. He reaches up and touches my clipboard while looking directly in my eyes. Sam then moves off toward the books and toys on the floor in the middle of the room.

We did our best to remain as unobtrusive as possible. Samuel's interest in the recorder and the clipboard indicates that our presence was associated with some observer effect. However, a systematic search of field notes and audio-recording transcripts revealed no incidents where our presence seriously disrupted the normal routines and activities of the storytimes.

Discussion and Conclusion

While labor intensive, participant observation clearly is an appropriate and effective method for studying babies and toddlers in public library settings. As can be seen in the data collected in our exploratory study of baby storytime, actions speak as loudly as words. Children's spontaneous, naturally occurring behavior may be observed and recorded and provides rich information about their interaction with library staff, materials, space, and services. However, in order for participant observation to be successful in group settings such as storytimes, researchers are advised to adopt the following practices:

- Remember that storytimes for babies and toddlers are complex and usually noisy events, with many activities of different types going on at the same time
- Use multiple observers to capture as much data as possible; however, do not use so many observers as to be intrusive
- Divide observation duties among the observers so that each is responsible for and can focus on a reasonable number (three to five pairs) of participants
- Expect the children to move around as soon as they are developmentally able to do so and develop strategies for tracking movements in your field notes
- When audio-recording library programs, use multiple recorders spread throughout the activity space in order to accommodate for noise blocking
- Recognize that getting informed consent requires flexibility and diligence due to the large number of participants, visitors, and new families; remember that special permission may be needed from ethics review boards and provisions made to avoid observing attendees before they have provided informed consent
- Ensure that child participants do not experience harm through the research process by continually looking carefully for signs of distress
- Protect the confidentiality of babies and toddlers by using pseudonyms in all research reports
- Remember that gender, ethnicity, and power all play a role in the relationship between adult observers and children and take care to minimize the impact of these variables as much as possible
- Whenever possible, select observers who are familiar with baby storytimes, share characteristics such as gender and motherhood with the adult caregivers, and are willing and able to dress and act like the adult participants so as to enhance their ability to gain and maintain access
- Systematically monitor data collection sessions and analyze transcripts and field notes for incidents of observer effect
- To increase the trustworthiness of your data and reduce observer effect through habituation, plan to observe at multiple sessions of the same storytime program series
- Use peer debriefing immediately after each program to confirm and help make sense of observations and to theoretically focus subsequent data collection

Appendix: Useful Resources

Guides to Observing Children

For the last sixty years instruction in the observation of young children has played an important part in the training of educators. Many guides

have been published, generally covering children from birth to age eight. While few of these are scholarly research method texts and most emphasize home, daycare nursery school, kindergarten, and early grade classroom settings, they all contain practical advice and many examples that demonstrate how to observe babies and toddlers and that are readily applicable to other settings such as libraries. Notable recent examples of these texts include the following.

- Bentzen, W. B. (2005). *Seeing young children: A guide to observing and recording behavior* (5th ed.). Clifton Park, NY: Thompson Delmar Learning.
- Billman, J., & Sherman, J. A. (2002). *Observation and participation in early childhood settings: A practicum guide* (2nd ed.). Boston, MA: Allyn and Bacon.
- Cohen, D. H., Stern, V., & Balaban, N. (1997). *Observing and recording the behavior of young children* (4th ed.). New York: Teachers College Press.
- Hobart, C., & Frankel, J. (2004). *A practical guide to child observation and assessment* (3rd ed.). Cheltenham, UK: Nelson Thornes.
- Sharman, C., Cross, W., & Vennis, D. (2004). *Observing children: A practical guide* (3rd ed.). London: Continuum.

Child Development Texts

Child development texts for early childhood educators provide good overviews of the physical, psychosocial, and cognitive development of babies and toddlers illustrated with real-life examples and case studies. Usually included is a chapter on methods for studying and observing very young children.

- Caulfield, R. A. (2001). *Infants and toddlers*. Upper Saddle River, NJ: Prentice Hall.
- Charlesworth, R. (2004). *Understanding child development* (6th ed.). Clifton Park, NY: Delmar Learning.
- Puckett, M. B., & Black, J. K. (2005). *The young child: Development from prebirth through age eight* (4th ed.). Upper Saddle River, NJ: Prentice-Hall.

Research Method Texts

The following books on doing research with children are exceptional in that they are well referenced, include examples from the research literature, and address the theoretical aspects of research design. Though these titles cover children from birth through about twelve years, each has a significant section on infancy and toddlerhood.

- Greig, A., & Taylor, J. (1999). *Doing research with children*. London: Sage.

Includes sections on theoretical approaches, appropriate research methods, and ethical considerations. The book is especially strong in encouraging an understanding of the unique nature of children as research participants.

- Fraser, S., Lewis, V., Ding, S., Kellet, M., & Robinson, C. (2004). *Doing research with children and young people.* London, UK: Sage.

In addition to separate chapters on the main stages of childhood and adolescence, this book has excellent coverage of issues such as power, ethics, gender, diversity, and involving children in the research process. The editors work in the fields of health and child development.

- Graue, M. E., & Walsh, D. J. (1998). *Studying children in context: Theories, methods, and ethics.* Thousand Oaks, CA: Sage.

An in-depth look at doing fieldwork with children, including how children have been conceptualized, the role of theory in research design, ethics, the role of the researcher, field research data collection methods, analysis, and report writing. Selected case studies illustrate concepts. Graue and Walsh work in the discipline of education.

- Pellegrini, A. D. (1996). *Observing children in their natural worlds: A methodological primer.* Mahwah, NJ: Lawrence Erlbaum.

Step-by-step instructions for designing, conducting, and analyzing diverse types of observation studies with children in field settings. Pellegrini comes from the discipline of education.

Notes

1. This research was funded by the American Library Association's Carroll Preston Baber Research Grant Award (The Young Child/Caregiver Storytime Program as Information Ground, 2003) and the Social Sciences and Humanities Research Council of Canada (Doing Early Learning: An Observation Study of Early Learning Programs, 2005–2008).
2. I wish to acknowledge and thank Dr. Pamela J. McKenzie (Assistant Professor, Faculty of Information and Media Studies, University of Western Ontario), co-investigator in both the Baber and SSHRC projects, and research assistants Brandi Borman and Kirsty Moffatt (MLIS students at the University of Western Ontario), whose work contributed to this article.

References

American Library Association (ALA), Association of Library Services to Children, Preschool Services and Parent Education Committee. (1990). *First steps to literacy: Library programs for parents, teachers and caregivers.* Chicago: American Library Association.

American Library Association, (ALA) Association of Library Services to Children. (1997). *Born to read: How to nurture a baby's love of learning.* Chicago: American Library Association.

Carey, R. F., McKechnie, L., & McKenzie, P. J. (2001). Gaining access to everyday life information seeking. *Library & Information Science Research, 23*(4), 319–334.

Caulfield, R. A. (2001). *Infants and toddlers.* Upper Saddle River, NJ: Prentice-Hall.

Christensen, P., & James, A. (2000). Introduction: Researching children and childhood: Cultures of communication. In P. Christensen & A. James (Eds.), *Research with children: Perspectives and practices* (pp. 1–8). London: Falmer Press.

Cohen, D. H., Stern, V., & Balaban, N. (1997). *Observing and recording the behavior of young children* (4th ed.). New York: Teachers College Press.

Dixon, J., & Dowd, F. S. (1993). Should public libraries program for babies? A national survey of children's coordinators. *Public Libraries, 32*(1), 29–36.

Dowd, F. S., & Dixon, J. (1996). Successful toddler storytimes based on child development principles. *Public Libraries, 35*(6), 374–380.

Feinberg, S., & Deer, K. (1995). *Running a parent/child workshop.* New York: Neal-Schuman.

Feinberg, S., Kuchner, J. F., & Feldman, S. (1998). *Learning environments for young children: Rethinking library spaces and services.* Chicago: American Library Association.
Fine, G. A., & Sandstrom, K. L. (1988). *Knowing children: Participant observation with minors.* Newbury Park, CA: Sage.
Flatow, K. M. (1997). Programs for babies in public libraries. *Illinois Libraries, 79*(3), 107–109.
Graue, M. E., & Walsh, D. J. (1998). *Studying children in context: Theories, methods, and ethics.* Thousand Oaks, CA: Sage.
Greene, E. (1991). *Books, babies and libraries: Serving infants, toddlers, their parents and caregivers.* Chicago: American Library Association.
Greig, A., & Taylor, J. (1999). *Doing research with children.* London: Sage.
Holmes, R. M. (1998). *Fieldwork with children.* Thousand Oaks, CA: Sage.
Langston, A., Abbott, L., Lewis, V., & Kellett, M. (2004). Early childhood. In S. Fraser, V. Lewis, D. Ding, M. Kellett, & C. Robinson (Eds.), *Doing research with children and young people* (pp.148–160). London: Sage.
Maddigan, B., & Drennan, S. (2003). *The big book of stories, songs, and sing-alongs: Programs for babies, toddlers, and families.* Westport, CN: Libraries Unlimited.
McKechnie, L., & McKenzie, P. J. (2004). *The young child/adult caregiver storytime program as information ground.* Paper presented at Library Research Seminar III, October, Kansas City, MO.
Nespeca, S. M. (1994). *Library programming for children with young families.* New York: Neal-Schuman.
Pattman, R., & Kehily, M. J. (2004). Gender. In S. Fraser, V. Lewis, S. Ding, M. Kellett, & C. Robinson (Eds.), *Doing research with children and young people* (pp. 131–144). London: Sage.
Robinson, C., & Kellett, M. (2004). Power. In S. Fraser, V. Lewis, S. Ding, M. Kellett, & C. Robinson (Eds.), *Doing research with children and young people* (pp. 81–96). London: Sage.
Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.
Thompson, R. A. (1992). Developmental changes in research risk and benefit: A changing calculus of concerns. In B. Stanley & J. F. Sieber, (Eds.), *Social research on children and adolescents: Ethical issues* (pp. 31–64). Newbury Park, CA: Sage.

Lynne McKechnie is an Associate Professor in the Faculty of Information and Media Studies at the University of Western Ontario where she teaches children's literature, library services for children, everyday life information practices, research methods, and reference. Dr. McKechnie has been named as the first Cleary Professor (children's literature and youth services) at the Information School at the University of Washington, a position she will fill in a visiting capacity until September 15, 2007. Dr. McKechnie's current research project, "Doing Early Learning: An Observation Study of Early Learning Programs," is supported by a standard research grant from the Social Sciences and Humanities Research Council of Canada. She is co-author (with Catherine Ross and Paulette M. Rothbauer) of *Reading Matters: What the Research Reveals about Reading, Libraries, and Community* (Libraries Unlimited, 2005) and co-editor (with Karen Fisher and Sanda Erdelez) of *Theories of Information Behavior* (Information Today, 2005). Dr. McKechnie is editor of the *Canadian Journal of Information and Library Science,* Coordinating Chair of the ALISE Research Committee, and Chair of the ASIST SIG Use Award Adjudication Committee.

Systematic Reviews and Librarians

K. Ann McKibbon

Abstract

Systematic reviews are review articles that are completed using predefined methods to minimize bias inherent to observational studies. Systematic reviews are important to librarians because they integrate evidence across studies or data resources to provide knowledge that is useful to good decision making in our profession. In addition, as more systematic reviews are being published in many disciplines, librarians are being asked to assist with the production of them—comprehensive searching is vital to the strength of the reviews. This article describes the process of producing systematic reviews and also describes their use. Librarians can acquire the skills necessary to use and produce high-quality systematic reviews.

Introduction

This article is designed to introduce librarians, both practitioners and researchers, to systematic reviews. I plan to set the context of this article by describing a scenario, define what a systematic review is in relation to all review articles, briefly discuss the history of systematic reviews, and list why they are important to librarians and why they are done. I will also describe the research strengths and quality indicators, show the steps in the production of a systematic review, discuss how one can find them across databases, and resolve the scenario. I include examples throughout the article from the disciplines of library and information science (LIS) as well as health care, the area of librarianship in which I have spent the past twenty years. The examples in this article are ones that I chose to show a specific aspect of systematic review production or use and represent a range of quality and content.

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 202–215
© 2006 The Board of Trustees, University of Illinois

Scenario

You have just been promoted to be the head of one of the smaller inner-city branches of your local public library. One of the reasons you got the job is that you stressed the need for evidence from sound research to back up your library's services and collections. Your branch has a long and strong tradition of many programs and is especially proud of their bibliotherapy programs run in conjunction with the local Department of Public Health. Your library director has just called and asked you to provide evidence concerning the effectiveness of the bibliotherapy programs. You sit down at your terminal and quickly find 252 articles in the National Library of Medicine's (NLM) PUBMED database. You sigh and wish that someone else besides you could "pull" all of these papers together and come up with a sound, evidence-based bottom line for bibliotherapy.

While you are thinking, the phone rings again and it is the director of the Department of Public Health. The city has just realized that their teen pregnancy rate is well above national and state levels and has started to push the Public Health people to "do something." The director, a long-time supporter of your services, asks if you could spare one of your librarians to help them search for and collect literature on prevention of teen pregnancy. The Health Department needs to write a report summarizing the evidence on the effectiveness of various approaches to preventing teen pregnancy and plan for new programs. Again you are faced with compiling the information on a certain topic or area and having it ready for others to apply—in other words, a systematic review of the literature.

What Is a Systematic Review?

The research world recognizes two sorts of review articles, both of which are important. Narrative reviews are opinion pieces done by an expert in the field. They are often broad based, written by a single author, and lack formal summaries of whole bodies of knowledge. Narrative reviews provide valuable coverage of an area of knowledge or an introduction to a topic, similar to what would be found in a textbook chapter. Two useful examples of narrative reviews in LIS include bibliotherapy and bullying (Gregory & Vessey, 2004) and public libraries and ethnic minority communities in the UK (Elliott, 1999). Both were written by experts who summarized the content area of their respective topics. Students and those interested in a general summary of a topic value these expert (narrative) reviews.

Systematic reviews, on the other hand, are often much more narrowly focused and are written by a team of researchers who represent a range of skills and interest in the topic. Cook, Mulrow, and Haynes (1997) define them as reviews that assemble, critically appraise or evaluate, and synthesize the results of primary studies in an integrative approach. They continue by listing the features of a well-done systematic review:

1. It defines the question to be addressed precisely and explicitly
2. It includes a replicable search strategy (for example, databases, terms, years, language restrictions, and other limits)
3. It uses pre-set inclusion and exclusion criteria to select articles or data sources that will be summarized in the review

To illustrate the features of a systematic review, it is worth looking at a study by Weightman and Williamson (2005). These authors wanted to examine the research on the value and impact of information about patient care provided by health sciences librarians. Their goal was "to review studies looking at the value and impact of library services on health outcomes for patients and time saved by health professionals" (p. 5).

They searched six databases (Education Resources Information Center [ERIC], PUBMED, Library and Information Science Abstracts [LISA], PREMDELINE, EMBASE©, and Cochrane Database of Systematic Reviews©) using twenty terms in various "and" and "or" combinations. They also searched Google, did a hand search of two journals, and checked all bibliographies of the articles they retrieved. In addition, they contacted authors, used personal reprint collections, consulted peers, and submitted emails to several discussion groups and listservs. Their inclusion and exclusion criteria for individual studies spanned several paragraphs. Starting with 320 papers, they reduced the total to 68 papers on the first screen for inclusion and, with closer examination and data extraction, reduced the papers to the final 28 studies that were analyzed in their systematic review.

Their systematic review is well done and provides a strong base for building and maintaining professional library services for patients and health care providers. The bottom line across twenty-eight studies is that professionally led library services do impact health outcomes for patients and save time for health care staff.

Systematic reviews can be qualitative or quantitative in nature. The former combines the information from the studies and describes results in a verbal format (for example, "four studies of clinical librarian projects suggested that professionals saved time"; "two studies showed evidence of cost-effectiveness"). The data from quantitative studies can, but not necessarily, be combined numerically and statistically. If this type of numerical and statistical combining is done (see, for example, Anderson et al., 2005, who looked at self-help books for depression), the systematic review is also classified as a meta-analysis.

History of Systematic Reviews

Systematic reviews date back to 1904 (Pearson, 1904). Although the first few were in medicine, many researchers and clinicians in education and psychology, as well as other social sciences disciplines, have done much to develop and improve systematic review methods and reporting. Much of

this early work was concentrated in the 1960s and 1970s (see, for example, Glass, 1976). Systematic reviews are becoming more common in the health sciences, where researchers have built upon the evidence-based practice movement. Many other disciplines are producing more systematic reviews, in part because of the ease of finding studies and data sources to combine as well as advancement of systematic review methods. For anyone interested in learning more about systematic reviews, a very readable work describing the systematic review process in health care is available on the Internet from the Millbank Memorial Fund (Moynihan, 2004). Many discipline-specific texts also exist on systematic review and meta-analysis production.

Why Are Systematic Reviews Important to Librarians and Librarianship?

Systematic reviews and meta-analyses are important to librarians for two main reasons. First, they help us build and make sense of our own research base. Using systematic reviews we can more easily identify our strengths while finding out where gaps exist. Using information from systematic reviews, we can implement and justify valuable services and programs while stopping or bypassing those programs that have not been shown to be beneficial. A well-done systematic review means that individuals do not have to collect and analyze primary studies for every decision they make.

Second, we are the professionals who have access to and who can effectively access the world's knowledge. Systematic reviews and meta-analyses and their results are only as strong as the evidence that is gathered for analysis. Librarians have been key players in many systematic reviews. The role includes locating published reviews and identifying and obtaining studies for new systematic reviews. Harris (2005), a health sciences librarian who has done considerable work in the field, summarizes the role of an information scientist in the systematic review process. She outlines many considerations for anyone interested in being a team member on a systematic review project and the roles that a professional librarian can take in the process. Because systematic reviews are important to librarians, I feel that it is important for us to know their strengths and weaknesses as well as understand the production process so that we can not only use them effectively but also assist in their production.

Why Are They Done?

Systematic reviews are done for many academic and application-based reasons. They are useful in the following cases.

- *Too much information is available.* For example, Ondrusek (2004) studied the attributes of research on end-user behavior for both online catalog and document retrieval systems—an almost insurmountable task. Her final report (45 pages) brings together 163 studies published in 175 ar-

ticles. Analyses were done on year of publication, research populations, research methodologies (qualitative and quantitative), performance analyses (outcomes and obstacles), factors affecting performance (tasks, systems, and end-user traits), and historical trends. She not only summarizes her findings in tables and narrative form but also talks about the implications of her findings for librarians who make decisions about search engines for Web sites. Anyone interested in end-user searching would save much time by reading her study or working from the bibliography.

- *Too little information is available.* Some issues occur so infrequently that to understand them fully and systematically means going to previous literature. Even case reports of single episodes can provide integration and knowledge. This type of systematic review of uncommon occurrences is more common in health care than in some other disciplines. For example, scoliosis (curvature of the spine that can sometimes be fatal) is thought to be a genetic disorder. To remove some of the environmental and other biases in research into the causes of scoliosis Kesling and Reinker (1997) sought to study twins with scoliosis. Rather than waiting for occurrences of twins with the disorder to happen, the authors went to the literature. They found published data on 100 cases of sets of twins with scoliosis at adolescence. Sufficient data for analysis were available for 68 sets of twins. Analysis showed that genetics does play a strong role in adolescent idiopathic scoliosis.
- *To resolve discrepancies.* Systematic reviews can often (but not always) help uncover the truth about hotly debated issues. A recent systematic review on the effectiveness vitamin C in preventing the common cold comes to the conclusion that despite thirty or more years of controversy, vitamin C does not protect against colds. The review carefully lays out the history of the issue of benefit from vitamin C, including publication of two systematic reviews that came to very different conclusions (Douglas et al., 2006).
- *To plan for new research.* Research need is one of the most important drivers of systematic review production. Established researchers or those with a well-defined research direction produce systematic reviews to justify and plan future work, build on the work of others, communicate their findings, and position their research ideas in the field. Reading systematic reviews produced by other researchers and practitioners, especially the background, conclusions, and discussion sections, can stimulate new ideas and projects for students and researchers seeking to modify or establish research programs.
- *To provide teaching/training materials.* Both narrative and systematic reviews are effective for teaching graduate students because they cover the research and general aspects of a specific topic in greater depth than one would find in a textbook chapter. Systematic reviews, by definition,

are produced using stronger methods and therefore are less prone to bias than narrative reviews. To promote the continued use of research material by their graduate students, educators should set the example by using systematic rather than narrative reviews. Hopefully, then, they will make decisions using evidence derived from studies using the strongest possible methods (evidence-based LIS).

In summary, many reasons exist for producing and using systematic reviews. Many more systematic reviews are being published across disciplines, and in some areas such as health care, education, and psychology researchers and practitioners rely heavily on them. I will now move to a discussion of the production of systematic reviews, the steps that must be followed, and how to search for and find published systematic reviews.

What Are the Strengths and Weaknesses of a Well-Done Systematic Review?

As with all research, high-quality systematic review articles must conform to accepted methods of production. In most research methodology classifications, systematic reviews are considered to be observational and retrospective. Therefore, they must conform to standard research methods common to all research projects as well as methods unique to observational studies.

First, systematic reviews must be preplanned. This involves development of a protocol that is based on a concise research question and lists the steps in production. The steps must be described in sufficient detail so that those involved in the process understand the tasks and the tasks are completed consistently, correctly, and efficiently. (The steps involved in doing a systematic review are discussed below.) A description of the question and the steps taken must be evident in the published report of the review process so that any reader can identify that careful preplanning was done and replicate the steps if necessary.

Bias is important in observational studies. It can be thought of as any factor, situation, or influence that, when acting alone or together, systematically distorts how we see or report data. Biases take us unknowingly away from the "truth" in research. To overcome or reduce bias, researchers use the strongest methodologies possible (for example, randomized controlled trials). Bias is also reduced by very careful execution of all aspects of the study. To counteract the potential for bias in systematic reviews, researchers who conduct them must emphasize the care they took to develop and carry out the entire process. In addition, they must report the process in detail in their published reports so that any external person can review what they did and even replicate the process to check outcomes.

After setting the question, the steps involved in a systematic review include the identification of potential studies or data sources, selection of

studies/sources, data extraction, combining and analyzing the data, and presentation of the findings. Each of these steps is expanded in the next section.

Steps in the Production of a Systematic Review

Systematic reviews take considerable time and resources to complete. Broad topics, such as Ondrusek's (2004) review of the research that evaluated end-user online searching behavior with its analysis and synthesis of 175 articles, would take about a year to complete. It is noteworthy that many graduate schools offer courses on systematic reviews and meta-analysis and estimate that it would take approximately 600 hours to complete a narrowly focused review using a team of two to five reviewers. Because of the time needed and skills involved (information retrieval, content expertise, and research methods experience), an interdisciplinary team often works together to produce systematic reviews. Any systematic review project starts with formulation of the question to be addressed.

Question Formulation

All good research is question driven. A well-formulated question for a LIS topic would likely include a description of who was involved (for example, library users, undergraduates, other libraries or librarians), what was being studied (for example, bibliotherapy for bullying, mother-toddler story programs, online instruction for health literacy), the outcomes in which one is interested (for example, increased use of the collection, higher computer literacy), and what studies or data to collect and combine (for example, surveys done by public libraries in cities of similar size to yours, evaluations of online training versus tutorials to increase use of your catalog, randomized controlled trials of giving books to young mothers at well-baby visits). An example of a comprehensive statement or aim from a systematic review relevant to health sciences librarianship is "to establish an evidence base for CL [clinical librarian] programmes . . . to determine, from the literature, whether CL programmes

1. are used by clinicians
2. have an effect on patient care
3. have an impact of clinicians' use of literature in practice
4. are cost-effective" (Winning & Beverley, 2003, p. 11).

All members of the team should work to develop and perfect the question because it guides the rest of the review production process. Formulation of the ideal question can take time. Once the question is complete, the review process moves on to identifying potential studies or data sources.

Searching for Studies (Information Retrieval)

The searching or retrieval step is where librarians who are members of or consultants to a systematic review team play a major role. The research

question will guide the search process by providing content terms to be translated into structured vocabulary, synonyms, and text words. In conjunction with team members, librarians will have to decide on the most relevant databases to search; years to be included; and limits based on methodology (for example, only randomized controlled trials), geography, language, and patient characteristics (for example, only adolescents). Comprehensive searching also can include hand searching of specific journal titles using the predefined criteria.

Searching is often done in two phases. In the first phase, the goal of the search is to identify published narrative and systematic reviews. If a relevant systematic review is already available, the project could end. If the identified review is on target but older, the research team can build upon the older review and choose not to include studies from it in the newer one, that is, produce an update rather than a complete review. If the reviews retrieved are not exactly on target, they can, at least, provide insight into search terms and database selection as well as potential citations for inclusion in the new review.

After searching for published reviews, the searching proceeds to identify potential original studies. These studies come from three main sources: primary searches in established databases and hand searches of specific journal titles; personal knowledge (team members' reprint files) and personal contact with peers and experts in the field; and "snowballing," whereby the team members find potential citations in bibliographies of reviews and original studies as well as perform citation tracking of important and older studies using resources such as Science Citation Index, Social Science Citation Index, and Arts and Humanities Citation Index. The database and hand-searching procedures are set before the study starts (preplanned) and the "snowball" accumulation occurs as the study progresses. Greenhalgh and Peacock (2005) showed that in 495 studies and systematic reviews of complex health care evidence, 30 percent of the data sources and articles were identified using protocol-based searching methods, 24 percent came from personal knowledge or peers, and 51 percent came from snowballing. Searching done for systematic reviews must be comprehensive and is often complex and iterative.

Comprehensive searching is the foundation of systematic reviews, and librarians are considered to be the experts in this area. I summarize the databases and searching performed by Winning and Beverley (2003) in their review of clinical librarianship. They used nine search phrases in a free-text and thesaurus approach with multiple truncations. They searched many databases in the following areas:

- Medicine (for example, PUBMED, EMBASE)
- Other health care disciplines (for example, Cumulated Index to Nursing and Allied Health Literature [CINAHL], British Nursing Index, Allied and Complementary Medicine Database [AMED], HealthSTAR)

- Science (for example, Science Citation Index)
- Social Science (for example, Social Science Citation Index, Applied Social Science Index and Abstracts [ASSIA])
- Information Science (for example, Library and Information Science Abstracts [LISA], Information Service for Physics, Electronics, and Computing [INSPEC])
- "Grey literature" (unpublished studies and sources) (for example, Health Management Information Consortium, Index to Theses, the National Research Register, Current Research in Britain, and COPAC—catalogs of twenty-four major UK universities plus the British Library, the National Library of Scotland, and the National Library of Wales)

Winning and Beverley (2003) also did citation tracking of identified studies, as well as checking bibliographies of studies and published review articles. Hand searching was done in the *Bulletin of the Medical Library Association* and *Health Information and Libraries Journal*. They contacted experts in the field to ask for other published and unpublished studies, an information-retrieval step often included in systematic reviews.

A high-quality systematic review includes a list of each database searched with all limits described, terms used, and other searching processes. For some reviews all of this information is in the published report, while for other reviews a link to a Web site or an invitation for email requests are included.

After the predefined searching is finished, citations are downloaded, combined into one list with duplicates removed, and sorted for easy screening by members of the team. The members often use titles, abstracts, and subject headings to do this initial screening. It is not unusual to have collections of several thousand citations for review using predefined inclusion and exclusion criteria.

Inclusion and Exclusion Criteria to Select Studies for Analysis

The study protocol developed before the searching started needs to include the selection criteria (inclusion and exclusion) that define which articles are to be included in the analyses. By predefining and adhering to the selection criteria, bias in choosing studies for inclusion is minimized. Furthermore, by publishing the criteria along with the search strategies and process, readers can verify that studies were chosen using methods that minimize bias and determine why other studies were excluded. A useful book chapter on systematic reviews (Egger, Dickersin, & Davey Smith, 2001) provides insight into decision making related to selecting studies, as well as a good discussion on publication bias, that is, the propensity of researchers and editors to publish studies that have "positive" results. Trials of "negative" results (those that show no benefit or results that are "disappointing" to the researcher) are published less often. If they are published, a considerable time lag can occur between when the study was done and when it can be found in print. Oftentimes, the journal is of lesser quality and may not

even be indexed in the major databases of that discipline. This publication bias leads to inclusion of a higher proportion of studies with positive results while ignoring trials with negative or disappointing findings.

Weightman and Williamson (2005) used the following inclusion and exclusion criteria in their systematic review of the value and impact of information provided through library services for patient care.

Inclusion Criteria

- Reports that included a formal evaluation using any research methodology
- Services studied were from professionally led libraries
- Services were provided to health professionals
- At least one outcome had to be provided that related to
  - Health benefits for patients, members of the public, or both
  - Time saved by the health professionals

Exclusion Criteria

- Library services were based only on virtual provision of established resources
- Studies of information skills training
- Specialist training to specific groups of health professionals (e.g., family physicians) outside the traditional library setting (p. 12)

Screening was done on 320 papers, and 28 were included in the published review. Both authors screened studies for inclusion; one author did the initial screening and the second author verified the results and resolved problems. Data checking and duplication of study selection and data extraction are effective methods of minimizing bias. After identifying the studies to be included in the review, each article is obtained in full text. The next step is to extract the data.

Data Extraction

Data extraction from each study or paper is the next step in the process. Using the protocol and its predefined data elements, the team develops a data extraction form to be used by the readers. Each paper is carefully read, often by two people with resolution of differences made through consensus or by bringing in a third party. Data forms in paper and increasingly in electronic format are used to ensure conformity and reproducibility. Some of these forms have multiple pages. Authors of systematic reviews may offer to provide copies of their data extraction forms to anyone who is interested in them. For anyone who is new to the systematic review process, collection of several of these extraction forms from reviews similar to the one you are working on can provide insights and templates.

At the same time as data extraction, individual studies or papers can be evaluated for the quality of their methods. Bias can occur in that lower-quality studies often inflate or exaggerate their findings; thus, analysis of all

studies may provide a different answer than analysis of only the high-quality studies. A well-done systematic review will provide data on the method of quality assessments of each study or data source as well as a description of the individual studies and their characteristics. (This quality evaluation may be more important in systematic reviews that are also meta-analyses.)

Data Analysis and Presentation

Data analysis proceeds after data collection. If the data across studies/papers/data sources can logically (and statistically) be analyzed to provide one final answer to the question, often in numerical form, the systematic review becomes a meta-analysis. Most of the reviews I discuss in this article are non-meta-analysis systematic reviews—the data on the studies are not numerically combined but presented more in a "vote-counting" manner. For example, Weightman and Williamson (2005) extracted the data and summarized what they were across the studies; no numerical combining took place. They stated: "The higher quality traditional library studies . . . suggest effects of impacts between 37 and 97% on general patient care, 10–31% on diagnosis, 20–51% on choice of tests, 27–45% on choice of therapy, and 10–19% on reduced length of stay" (p. 4).

These data on improved care and patient outcomes from studies show the worth of health sciences librarianship and are impressive. Systematic reviews such as this one set the standard for other branches of librarianship to provide evidence of their worth using systematic review techniques. Not everyone has the resources to be able to complete a high-quality systematic review. Many of them, however, have been done and can easily be found, especially by librarians with good search skills.

How Do I Find Systematic Reviews?

Most of the electronic databases include systematic reviews. The examples I have used in this article came from NLM's PUBMED and LISA rather than personal files. In LISA I used variations on the terms systematic review(s), systematic overview(s), and meta-analys(e/i)s as well as meta-nalysis and metaanalysis. For the health-related databases, several hedges (predetermined search strategies) exist. For example, the University of York in the United Kingdom maintains a database of these hedges for systematic reviews and meta-analyses (http://www.york.ac.uk/inst/crd/search.htm). In addition, the Cochrane Collaboration, a volunteer organization of health professionals (http://www.cochrane.org/index.htm), collects randomized controlled trials in all areas of health care and uses them to publish systematic reviews. Librarians have been involved with the Cochrane Collaboration since its inception in the mid-1990s. The Cochrane Library has over 1,000 systematic reviews, including several of interest to librarians. The library also includes resources to help those who want to learn more about systematic reviews and their production. The Campbell Collaboration

(http://www.campbellcollaboration.org/index.html) is a similar volunteer organization whose mandate includes collection of studies and production of systematic reviews in areas of education and social and behavioral sciences. Some of their systematic reviews are relevant to librarians (for example, impacts of after-school programs on student outcomes). The Campbell Collaboration provides opportunities to learn about systematic reviews and to publish library-related reviews. The Database of Abstracts of Reviews of Effects (DARE; http://www.york.ac.uk/inst/crd/crddatabases.htm) provides abstracts from systematic reviews as well as training material for those interested in systematic reviews and meta-analyses in health care, which is defined very broadly. DARE is produced by the Centre for Reviews and Dissemination at the University of York. Librarians have been involved with DARE since its inception. Access is free and efficient; it is a good place to look for published systematic reviews.

What Sort of Material Can Go Into Systematic Reviews?

Many systematic reviews exist, and the studies or data sources that can be effectively integrated to produce new knowledge are almost limitless. The majority of the systematic reviews I have discussed have selected and analyzed quantitative studies. Qualitative studies can also be synthesized using the techniques listed above. An excellent example of a qualitative systematic review summarizes parental attitudes toward childhood vaccination (Mills, Jadad, Ross, & Wilson, 2005). Worries about adverse effects and pain are major considerations for parents. With this review, those who provide immunization can better meet parental information needs and develop effective marketing methods.

A number of examples related to health care show the variety of material that can be integrated, such as an analysis of twenty published definitions of drowning (Papa, Hoefle, & Idris, 2005), and portraits from the fourteenth to the twentieth century to assess disease frequencies (Als et al., 2002). Systematic reviews can also be integrated into other systematic reviews when an abundance of information is present. This type of review is called a "meta-meta-analyses," or a systematic review of systematic reviews (Katerndahl & Lawler, 1999).

Resolution of the Scenario

At the start of this article we looked at a scenario where you, as the new director of a branch library, had received two important requests that could be addressed by either the production of new systematic reviews or identification of existing ones. You did some searching across health and psychology databases and found at least six high-quality systematic reviews that address the question of the effectiveness of bibliotherapy programs for a wide variety of conditions. The library director was very impressed with

your searching skills and greatly appreciated these reviews that reduced his/her workload.

You also found an incredibly detailed systematic review that outlines evidence (randomized controlled trials) of interventions to reduce unintended teen pregnancies (DiCenso, Guyatt, Wilan, & Griffith, 2002). The Public Health director was delighted with the review, in spite of the fact that the evaluated interventions (26 trials in 22 studies) did not show reductions in the rate of intercourse or pregnancy or improvement in the use of contraception. The Public Health director decided that the review was sufficient and no further evidence needed to be gathered or produced. You then decide that you will sign up for a systematic review course at your local university that fall, knowing that knowledge and experience with systematic reviews would be good for your career and job.

Summary

Systematic reviews are an important research method for librarians. These reviews are designed to collect evidence on a given topic from multiple sources using recognized and strong methods to minimize bias. By combining data and information from the collection of varied sources, established information is summarized and integrated and new information is obtained. By applying standard methods of research to avoid or minimize bias in data collection and analysis, we can advance knowledge and improve our services and programs. Librarians need to use and understand systematic reviews both inside librarianship and as partners in the production of high-quality systematic reviews in other disciplines. We already have many of the skills needed, and with some practice and training, we can become effective producers and users of systematic reviews.

References

Als, C., Stussi, Y., Boschung, U., Trohier, U., & Waber, J. H. (2002). Visible signs of illness from the 14th to 20th century: Systematic review of portraits. *BMJ, 325*(7378), 1499.

Anderson, L., Lewis, G., Araya, R., Elgie, R., Harrison, G., Proudfoot, J., et al. (2005). Self-help books for depression: How can practitioners and patients make the right choice? *British Journal of General Practice, 55*(54), 387–392.

Cook, D. J., Mulrow, C. D., & Haynes, R. B. (1997). Systematic reviews: Synthesis of best evidence for clinical decision. *Annals of Internal Medicine, 126*(5), 376–380.

DiCenso, A., Guyatt, G., Wilan, A., & Griffith, L. (2002). Interventions to reduce unintended pregnancies among adolescents: Systematic review of randomised controlled trials. *BMJ, 324,* 1426–1434.

Douglas, R. M., Hemila, H., Chalker, E., & D'Souza, R. R. D., & Tracy, B. (2006). Vitamin C for preventing and treating the common cold. *Cochrane Database of Systematic Reviews, 4,* CD000980.

Egger, M., Dickersin, K., & Davey Smith, G. (2001). Problems and limitations in conducting systematic reviews. In M. Egger, G. Davey Smith, & D. G. Altman (Eds.), *Systematic reviews in health care: Meta-analysis in context* (2nd ed., pp. 43–68). London: BMJ Books.

Elliott, P. (1999). Research into public libraries and ethnic minority communities in the UK, from Clough and Quarmby to Roach and Morrison: A review article. *Journal of Librarianship & Information Sciences, 31*(3), 174–177.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(1), 3–8.

Greenhalgh, T., & Peacock, R. (2005). Effectiveness and efficacy of search methods in systematic reviews of complex evidence: Evidence audit of primary sources. *BMJ, 331*(7524), 1064–1065.

Gregory, K. E., & Vessey, J. A. (2004). Bibliotherapy: A strategy to help students with bullying. *Journal of School Nursing, 20*(3), 127–133.

Harris, M. (2005). The librarian's roles in the systematic review process: A case study. *Journal of the Medical Library Association, 93*(1), 81–87.

Katerndahl, D. A., & Lawler, W. R. (1999). Variability in meta-analytic results concerning the value of cholesterol reduction in coronary heart disease: A meta-meta-analysis. *American Journal of Epidemiology, 149,* 429–441.

Kesling, K. L., & Reinker, K. A. (1997). Scoliosis in twins: A meta-analysis of the literature and a report of six cases. *Spine, 22*(17), 2009–2014.

Mills, E., Jadad, A. R., Ross, C., & Wilson, K. (2005). Systematic review of qualitative studies exploring parental beliefs and attitudes toward childhood vaccination identifies common barriers to vaccination. *Journal of Clinical Epidemiology, 58,* 1081–1088.

Moynihan, R. (2004). *Evaluating health services: A reporter covers the science of research synthesis.* Retrieved November 2, 2005, from http://www.milbank.org/reports/2004Moynihan/Moynihan.pdf.

Ondrusek, A. L. (2004). The attributes of research on end-user online searching behavior: A retrospective review and analysis. *Library & Information Research, 26,* 221–265.

Papa, L., Hoefle, R., & Idris, A. (2005). Systematic review of definitions of drowning incidents. *Resuscitation, 65,* 255–264.

Pearson, K. (1904). Report on certain enteric fever inoculation statistics. *British Medical Journal, 3,* 1243–1246.

Weightman, A. L., & Williamson, J. (2005). The value and impact of information provided through library services for patient care: A systematic review. *Health Information & Libraries Journal, 22*(1), 4–25.

Winning, M. A., & Beverley, C. A. (2003). Clinical librarianship: A systematic review of the literature. *Health Information & Libraries Journal, 20*(Suppl.1), 10–21.

Ann McKibbon is Associate Professor in the Department of Clinical Epidemiology and Biostatistics, Faculty of Health Sciences, at McMaster University. She has worked as a science and health sciences librarian in academic and special libraries and as a librarian researcher in a medical school for thirty years. She received her Ph.D. in medical informatics in 2005 by studying how risk and uncertainty in primary care physicians affects use of electronic information resources. She has taught evidence-based health care to librarians and health professionals over the past twenty years. She has also produced information resources for physicians and nurses and has done research into the production of hedges to identify clinically relevant and important studies in medical and other health databases. She is currently interested in how health professionals who use electronic information resources obtain incorrect answers to their questions. She has published two books and fifty peer-reviewed journal articles, which include several systematic reviews.