ARTICLES
Lisa A. Wishard and Linda R. Musser 131 Preservation Strategies for Geoscience Literature: New Technologies for an Old Literature
Karen M. Drabenstott, Schelle Simcox, and Eileen G. Fenton 140 End-User Understanding of Subject Headings in Library Catalogs

NOTES ON OPERATIONS
Barbara Nichols Randall 161 Spelling Errors in the Database: Shadow or Substance?
L. Suzanne Kellerman 170 Combating Whole-Book Deterioration: The Rebinding & Mass Deacidification Program at the Penn State University Libraries

FEATURES
Margaret Rohdy, Editor 177 Index to Advertisers
Nancy Williamson 178 Book Reviews
Nancy Williamson 186 In Memoriam: Phyllis Allen Richmond

43(3): 129–188
ISSN 0024-2527
If you have been searching for an easy way to authority control your library’s current cataloging, try LTI’s Authority Express service.

With Authority Express, a library uses the Internet to transmit a file of newly cataloged bibliographic records to LTI (via ftp). LTI immediately processes the records through its state-of-the-art authority control system. Then, at the library’s convenience, it logs into LTI’s ftp server to retrieve fully authorized catalog records, along with linked LC name and subject authority records.

**Authority Express**
- Keeps authority control current at an affordable price
- Integrates easily into existing workflows
- Lowers cost by reducing staff time spent on catalog maintenance
- Provides 1 hour turn around for up to 1,000 catalog records

*“Authority Control for the 21st Century”*

**Library Technologies, Inc.**
2300 Computer Avenue, Suite D-19 Willow Grove, PA 19090
(215) 830-9320  Fax: (215) 830-9422
(800) 795-9504  email: LTI@LibraryTech.Com

Visit our website at:
www.LibraryTech.Com
ARTICLES
Lisa A. Wishard and Linda R. Musser 131 Preservation Strategies for Geoscience Literature: New Technologies for an Old Literature
Karen M. Drabenstott, Schelle Simcox, and Eileen G. Fenton 140 End-User Understanding of Subject Headings in Library Catalogs

NOTES ON OPERATIONS
Barbara Nichols Randall 161 Spelling Errors in the Database: Shadow or Substance?
L. Suzanne Kellerman 170 Combating Whole-Book Deterioration: The Rebinding & Mass Deacidification Program at the Penn State University Libraries

FEATURES
Margaret Rohdy, Editor 177 Index to Advertisers
Nancy Williamson 178 Book Reviews
186 In Memoriam: Phyllis Allen Richmond

43(3): 129–188 ISSN 0024-2527
American Library Association
Association for Library Collections & Technical Services

For current news and reports on ALCTS activities, see the new ALCTS Newsletter Online at http://www.ala.org/alcts/alcts_news.
Preservation Strategies for Geoscience Literature: New Technologies for an Old Literature

Lisa A. Wishard and Linda R. Musser

In this article we describe the complex nature and history of the literature of the geosciences. We evaluate preservation technologies applicable to the formats found in the literature, such as oversize, color, and folded maps and plates. Technologies profiled include deacidification, encapsulation, storage options, and facsimile creation including xerographic, photographic, film, and digital reproduction. We present applications of the technologies and discuss suggestions for developing preservation priorities for the literature.

Researchers in the geosciences use a wide variety of information sources. While many of these sources are traditional library materials such as journal articles, books, and conference proceedings, geoscientists rely heavily on maps, technical reports, photographs, theses, and dissertations. Consequently, geoscience literature contains complex formats, often combined in a single volume. The frequent presence of folded pages, plates, oversize maps, and illustrations—many in color—makes geoscience materials difficult to preserve using traditional methods. An additional challenge is the value and heavy use of older material in the geosciences. The age of material used in this field is frequently older than in many other fields of scientific research, with materials more than one hundred years old still routinely in use. The complexity of formats found in the geosciences correspond to the literature of other disciplines, such as art, architecture, and engineering that also frequently use oversize, color illustrations (Commission on Preservation and Access 1992).

In addition to print resources, geoscientists also derive essential information from materials such as geologic specimens, drill cores, well logs, and digital data. North American geoscientists, recognizing the value of these original specimens, have begun an initiative under the auspices of the American Geological Institute to preserve these materials through the National Geoscience Data Repository (American Geological Institute 1998). Similar initiatives to preserve geologic specimens are underway in other countries as well.

 Unfortunately, preservation of the traditional print records of geoscience

LISA A. WISHARD (lawisha@sandia.gov) is senior technical information specialist, Technical Library, Sandia National Laboratories, Albuquerque, and LINDA R. MUSSER (lrm4@psu.edu) is head, Earth and Mineral Sciences Library, Pennsylvania State University, University Park. Manuscript received December 7, 1998; accepted for publication February 5, 1999.
knowledge has not received the same attention from the library community. In the U.S., national libraries such as the National Agricultural Library in agricultural sciences and the National Library of Medicine in the medical sciences have taken the lead in preserving the literature of several scientific fields. The U.S. has no national library devoted to the geosciences, however, and the U.S. Geological Survey Library, viewed by many as the de facto national geosciences library, has no established preservation program (Wishard 1999; Klimley 1984).

In this paper, we will look at the physical characteristics of geological literature, review past practices in preserving the geological literature, and examine advances in preservation technologies that are being applied to geoscience materials today. We will also profile several preservation projects in the geosciences and, finally, discuss the preservation initiatives of the Geoscience Information Society.

**Development of Geoscience Literature in North America**

Prior to 1823, geological research in the Americas appeared primarily in journal articles and monographs. In 1823, the first geological survey was conducted in North Carolina. Over the next 77 years, nearly 40 states, territories, and provinces followed suit so that, by the end of the century, the state and provincial geological survey had become a permanent part of many local governments. The 1800s saw increases in the amount of geoscience publications—not only in serials but also in the reports and volumes published by geological surveys (Stoffer 1988; Hazen 1980; Meadows 1974). By the 1850s, there were more than 50,000 pages published in the geoscience literature, primarily in serial publications (Menard 1971). During the next century, however, there was even greater growth in the number of publications in the field, with more than one million pages in print by 1950 (Menard 1971). Significantly, the majority of the publications published during this period were geologic surveys, which by the mid-nineteenth century outnumbered serial publications (Menard 1971).

The primary purpose that geologists had in the early period was to promote economic development through identification of useful mineral deposits. Much of their work remains relevant today and is routinely used by practicing geologists and engineers. Additionally, these early publications contain information on the growth and commerce of towns and cities, development of industries, details of Native American life and of other residents, and descriptions of local flora and fauna. Not surprisingly, today's users of the early geological literature of North America include not only geologists and engineers but also genealogists, historians, geographers, biologists, anthropologists, and others.

These early publications contain a variety of formats—text, maps (both color and black and white), photographs, and color illustrations. Scott (1993, 83) stated that geology was "one of the earliest sciences to employ color to encode information. Text cannot adequately describe, for example, the detailed and often complex subsurface relationships of stratigraphic formations without complementary illustrations." Klimley (1984) considered the heavily and variously illustrated geological publications of the nineteenth and twentieth century to be a worst-case scenario for application of preservation techniques. In addition, between 1850 and 1950—the era of highly acidic paper use—there was huge growth in the number of geological science publications. This is readily apparent upon examination of many of these documents, which exhibit the telltale signs of acidic paper—brittle, yellowing pages. Thus we have the current situation, in which early geosciences literature is increasingly useful to a broad range of users while the material itself is increasingly fragile and difficult to preserve.

**Historic Geoscience Literature Preservation Techniques**

To date, traditional preservation techniques have been applied to geoscience literature with some success. For exam-
ple, the collection of Sir William Logan, first director of the Geological Survey of Canada, received treatment on a volume-by-volume basis using deacidification and encapsulation in the collections of the Geological Survey of Canada (Kamel 1998). In many libraries, deteriorating materials have been rebound or placed in storage boxes. In other cases, folded maps have been unfolded and stored flat in attempts to reduce wear and breakage along fold lines.

Microfilming has been used as a preservation technique since the mid-1800s (Luther 1959), but difficulties related to use of color and the ability to film oversize images is still considered problematic for the geosciences. Most geoscience publications that have been microfilmed to date have been captured in black and white, resulting in the loss of the full integrity of many of the images. Examples of collections preserved in this way include the UMI collection of dissertations, theses, and miscellaneous geoscience serials, the American Natural History 1769-1865 microform set by Primary Source Media, and the Landmarks of Science series by Newsbank. The lack of color in this preservation process is extremely serious as most geologic maps use color to differentiate geologic formations from the underlying details of the base map. Scott’s (1993) evaluation of UCI-Ageology theses and dissertations found that 52% of the maps and plates evaluated were in color and used color to encode information. Additionally, in these collections, oversize maps were generally microfilmed in pieces so that only portions of the map can be viewed at a time, thus obliterating the overall characteristics or trends depicted in the original piece (Klimley 1984).

The preservation efforts undertaken thus far have affected only a tiny fraction of the geological literature requiring attention. Therefore, while there have been notable efforts to date, the bulk of preservation work of the geological literature lies ahead.

**Current Preservation Strategies**

The advent of improved technologies for color copying and new digitizing technologies provides methods for the reformatting and replacement of color and large size maps and illustrations. These, in conjunction with well-established preservation technologies such as deacidification and encapsulation, provide librarians with a wide range of tools necessary to preserve the geological literature. We will describe each of these technologies in reference to its use in preserving the geological literature. First, we will describe methods that focus on the preservation of the original document. Then we will discuss methods that preserve the content of the document.

**Deacidification**

Paper that is safely buffered and housed in a proper environment can last hundreds of years. The goal in deacidification is to absorb or reduce the acidic levels in existing paper publications to permit a longer shelf life. The process halts further acid damage but does not repair or improve the extant condition of the item. Deacidification does not remove the acid from the paper, nor does it strengthen brittle paper. In the process, the acid in the paper is chemically neutralized, which provides a chemical reserve that serves as a buffer for the paper fibers against future acid attacks. Deacidification can be a viable volume-by-volume treatment alternative for geoscience material.

Mass deacidification is the process by which batches of volumes are treated in a large treatment chamber. Several methods have been developed. In one of the more commonly used processes, volumes are immersed in a bath of alkaline submicron magnesium oxide particles for approximately 25 minutes. While in the bath, the volumes are agitated to increase contact of the pages with the magnesium oxide. Volumes are then dried in an evaporator chamber for several hours (Domach 1997; Dalrymple 1997). The use of mass deacidification has not been widely applied due to the perceived complexity of treating the multiformat geological literature.

In several pilot applications to geological collections, however, researchers suggest that mass deacidification is a viable
preservation option for this literature. Wick (1993) evaluated a 1992 mass deacidification process applied to separate maps, theses, and serial volumes with folded maps from the Harvard Kummel Geological Sciences Library collection and judged it to be satisfactory, with most volumes being successfully treated. Wick noted that there were some instances of color shifting on photographic plates and on folded maps within serial volumes, although colored plates and flat maps (including those that were hand colored) exhibited no shifts in color and were judged satisfactory. Gaydos (1998) found that during a recent trial at Pennsylvania State University, deacidification was successfully applied to volumes of clay-coated paper. Treated volumes were found to have an acceptable level of alkaline reserves. In a test of the volumes, Preservation Technologies, L.P., found alkaline reserves of 1.9%, in comparison with reserves of 0.6% found on similarly treated books from the Library of Congress.

The mass deacidification process provides a tool by which collections can be preserved in large quantities. Extra deliberation should still be applied, however, when brittle volumes with folded and photographic plates—items prevalent in collections of geologic material—are selected for mass deacidification. It should also be noted that acidic paper, even after deacidification, will be weaker than nonacidic paper and might require careful handling or reinforcement, such as binding or encapsulation.

ENCAPSULATION

Encapsulation is a commonly used technique to preserve sheet materials such as maps and illustrations. The sheet is placed between two pieces of polyester mylar. The edges are then sealed with an ultrasonic weld or with acid-free, double-sided tape. Gaps are left in the corners of the seal so that the piece can breathe. The static charge between the two sheets of mylar keeps the piece stationary and will actually hold together items that are in multiple pieces. Caution should be used when encapsulating hand-colored items, however, because "particles of color can be removed by the electrostatic charge which is created between the two sheets of mylar" (Scott 1993, 84). Mylar is an inert, inflammable material that has been subjected to many rapid-aging tests and has a life expectancy of several hundred years. The mylar enclosure reinforces the encapsulated piece so that even previously brittle material can be used with ease. Encapsulation is used to preserve both acidic and acid-free print materials. In some cases, for single-sided materials, a buffered alkaline paper backing can be inserted behind the piece in order to absorb some of the acid from the encapsulated item.

Encapsulation is one of the primary methods recommended for preservation of maps and oversize plates in their original form. Maps and plates that have been folded in a text block must be unfolded prior to encapsulation. Once encapsulated, maps should not be folded or re-folded. Encapsulation is a very effective method for preserving an original, though it can limit the use of a piece. For instance, once a map is encapsulated it becomes quite cumbersome to take into the field for active use—a common practice in the geosciences. In addition, encapsulation can also create problems when encapsulated items are separated from their explanatory and other accompanying materials.

Encapsulation has been used in many collections. Allen (1990) describes the New York Historic Map Preservation Project, where nearly 2,000 maps—primarily transportation and topographic maps on a variety of media, including wall, rolled, blueprint, and other various printed maps—were cleaned, deacidified, and encapsulated. Newman (1987; 1998) outlines how thesis and dissertation maps are cleaned and encapsulated at the DeLaMare Library of the University of Nevada, Reno. In an example of encapsulation and binding, staff at Pennsylvania State University Libraries deacidified, encapsulated, and post-bound 29 volumes related to Pennsylvania from the Geologic Atlas of the United States (Wishard 1999).
ENVIRONMENTAL AND STORAGE OPTIONS

The physical environment in which a piece is stored can also affect its longevity. Often out of the direct control of facilities managers, and not seen as glamorous to discuss (Wick 1991), collections are often imperiled by their very surroundings—by leaking labs, temperature and humidity fluctuations, etc. In addition to protecting against the physical climate of collections, storage of items in acid-free housings such as phase boxes and folders can help to diminish the cumulative deterioration of acidic paper. Certainly, properly sized shelving and map cases are important to ensure the longevity of materials. Archival storage of less frequently used and unique materials in a climate controlled facility can also help to preserve material.

FACSIMILE CREATION

The creation of facsimiles is a time-honored method for preserving the intellectual content of publications. Facsimile production is necessary to maintain access to information that might be stored in an inaccessible original medium. Until recent years, the primary methods available for facsimile creation were black and white photocopy reproductions or reproductions from monochromatic film. Photographic reproduction was also an option, albeit an expensive one. New technologies such as color photocopying, color microfilming, and digital scanning now offer additional possibilities for facsimile creation. These new facsimile technologies—which preserve extant volumes of varied mediums such as text, maps, and images—allow researchers to work with the full intellectual content of a piece as well as see the context of the material in its original format (Klimley 1984 and 1995; Gertz 1996).

XEROGRAPHIC AND PHOTOGRAPHIC FACSIMILES

Xerographic reproduction has been used successfully for many years to create facsimiles. Baird (1997) described Project Brittle, an example of a cooperative effort to use preservation photocopy technology to replace worn original material. The only limitation of this technique has been the inability to reproduce color. Color photocopy technology has advanced in recent years, however, such that inexpensive reproductions of standard-sized color pieces can be successfully produced that can allow the possibility of piecing together oversize items (Newman 1998; 1989).

Photographic facsimiles are extremely effective for reproducing details and colors of original documents, especially in oversize images. This technique’s primary drawback is its expense. While a photographic negative can be created at a reasonable cost, a print of an oversize image can be quite expensive.

FILM FACSIMILES

Microfilm remains the primary archival medium for text-based materials. Microfilm stored in ideal conditions can survive for half a millennium or more and would still be accessible with only candlelight and good eyesight, compared to the short usable life span of only 5 to 10 years for digitally stored materials (Conway 1996). However, producing high quality film images of oversize material still poses difficulties. If a large image can be captured in a single frame, the fine details such as roads, place names, and legends are often unreadable. Similarly, images sectioned onto several frames are often deemed unusable (Gertz 1996). The use of color in microfilming has, however, been successfully addressed (Klimley 1993; Newman 1989), despite the fact that capturing oversize images on single frames still proves problematic. Color microfilm technology has been successfully applied to color oversize images at Columbia University as part of the New York State Bulletin project (Gertz 1997) and at the Hargett Library at the University of Georgia (Henneberger 1994), where film images were used as a transition medium for creating digital copies of oversize images.

DIGITAL FACSIMILES

Developments in digital reproduction technologies have broadened facsimile creation options. Today, digital cameras and scanners allow the full integrity of the geologic original, including color, to be reproduced. The digital images can then be printed with color printers, plotters,
and photocopiers with nearly indiscernible differences in color from the original. In fact, the Commission on Preservation and Access (1992, 12) stated that there are "striking examples of visual material that has been digitized and subsequently reproduced in paper copy, sometimes clearer and more legible than the original document from which the digital copy was made."

The ability to produce digital facsimiles has increased the options for preserving the oversize media prevalent in the geological literature. These developments allow librarians to make preservation decisions based on the condition and content of material rather than by the limitations of traditional preservation options (Gertz 1996). These new digital options are especially relevant in relation to geologic maps and other oversize plates that have used color to encode the complex information that they depict.

Digital copies of materials can be stored as bitmapped images on a variety of magnetic and optical storage devices such as diskette, tape, or CD-ROM. Contemporary color technologies are well suited for reproducing geologic images on screen as well as in print. Digital reproductions of geologic images are viable at much lower resolutions than in many other fields (Gertz 1996; Commission on Preservation and Access 1992) because there typically only needs to be a discernible difference in color rather than reproduction of the fine color details. Satisfactory results have been achieved on pilot projects at Pennsylvania State University and other institutions (Allen 1998) using 24-bit true-color matching at 300 dots per inch (dpi) resolution as well as with 256 no-color matching at the same resolution. Gertz (1996; 1997) and Klimley (1993) examined the quality of digital images of the New York State Museum Bulletin and found acceptable resolutions at less than 300 dpi for oversize images. Allen (1998) profiled additional digital projects that suggest similar results, though he stated, "These efforts...have as yet led to no consensus as to whether it is really necessary to scan master images at a higher resolution than 300 dpi in 24-bit color, and no agreed upon minimum resolution for such archival images appears to be emerging."

Geological material has been digitally reproduced with a variety of techniques at different institutions. Heiser (1998) reported on a project at Indiana University in which early survey material for the state was scanned and stored as text files. Trombatore (1998) described a project in which documents at the University of Texas were scanned and stored as digital images of the actual page. At Washington University, a 1928 thesis was digitally reproduced by rekeying the text and scanning the images (McLeod 1998). Digital reproduction of oversize images has been explored at the Library of Congress (Fitzpatrick 1997), the United States Geological Survey (U.S. Geological Survey 1998), and the Virginia Board of Public Works (Library of Virginia 1998), among others.

Preserving material in a digital format is an attractive option because the captured image can be made accessible to a wider audience. However, the need of geoscientists to be able to take the material into the field can not be forgotten, so mechanisms for delivery of hard copy reproductions should be incorporated into any program to digitally preserve geologic material. In addition, the longevity issues raised by Conway (1996) remain problematic, as does the repeated and systematic refreshing (Commission on Preservation and Access 1992) that this transitory nature requires. The challenge of preserving information in electronic format cannot be overlooked. While digital technologies offer exciting new options for preservation, they pose new challenges for long-term retention and storage. Discussion of these challenges fall beyond the scope of this paper, however, and are covered well elsewhere in the literature (Allen 1998; Conway 1996; Klimley 1995; Rothenberg 1999).

**Priorities for Preservation**

Many of the projects profiled here are pioneering in scope and focus and in the use and development of new preservation methods. This work raises awareness of
the needs of preservation for the varied formats of information found within the geoscience field. In many of these projects, the accessibility and breadth of preservation techniques available to collection managers has also been demonstrated.

Such experiments in using these new technologies bode well for the future preservation of the geological literature. Indeed, one of the primary challenges facing geoscience librarians is not only to become educated regarding these technologies but also to raise awareness among their colleagues and the producers of the information that it is, at long last, possible to preserve this literature in all its facets—color, oversize, plates, etc.

In order to facilitate this process, the Preservation Committee of the Geoscience Information Society (GIS) (1997) summarized the challenges of preserving this literature and called for the development of a systematic international initiative to preserve the geoscience literature. DeFelice (1990) surveyed 107 librarians and found that 51% of the respondents indicated a willingness to participate in a cooperative preservation project. Wishard (1999) surveyed librarians at 489 libraries and found that less than half of the respondents had active preservation programs, and even fewer had preservation officers at their institutions. DeFelice found that there was no overall consensus on the question of the type of material that would need or will need preservation the most. Journals and folded maps were ranked highest, followed closely by books, flat maps, U.S. Geological Survey documents and theses, and thesis maps. Wick (1991) found that more than 70% of the respondents indicated some geographic designation—such as regionally, by country or state—to define how they choose candidates for preservation. Wishard (1999) found that when respondents were asked to identify geoscience literature most at risk and in need of preservation treatment, they ranked historic geologic maps as the highest priority for preservation treatment, followed by pre-1900 material. When asked to identify a focus for a preservation plan, most respondents indicated a preference for a geographic emphasis.

The continuing activities of the GIS Preservation Committee include coordinating preservation projects and acting as a clearinghouse that documents projects where geoscience literature has been preserved. In addition, the committee hopes to continue to demonstrate capabilities and raise awareness regarding the possibilities of preservation technologies for geoscience material.

CONCLUSION

Ultimately, however, the preservation of the geological literature rests in the hands of the many librarians around the world working with this literature as well as those producing the literature. Publishers need to develop standards for future geoscience publications to use preservationally sound material such as acid-free paper and digital formats that can be easily migrated, and apply such basic principles as limiting the size of maps and plates in theses and other publications (Newman 1998). Similarly, librarians need to develop priorities for preserving the material in their geoscience collections. Geoscience librarians need to move beyond the volume-by-volume approach to preserving their collections and realize that there are many viable and sophisticated means to preserve and conserve the literature of the geosciences. Publishers and librarians must work together to establish standards and priorities in order to ensure the survival of the geoscientific literature for future generations.

WORKS CITED


——. 1998. Thesis preservation project at the Mines/DeLaMare Library, University


End-User Understanding of Subject Headings in Library Catalogs

Karen M. Drabenstott, Schelle Simcox, and Eileen G. Fenton

In this article, we report on the first large-scale study of end-user understanding of subject headings. Our objectives were to determine the extent to which children and adults understood subdivided subject headings and to suggest improvements for improving understanding of subject headings. The 1991 Library of Congress Subject Subdivisions Conference suggested standardizing the order of subject subdivisions for the purpose of simplifying subject cataloging, which served as the impetus for the study. We demonstrated that adults understood subject headings better than children; however, both adults and children assigned correct meanings to less than half of the subject headings they examined. Neither subject heading context nor subdivision order had an effect on understanding. Based on our findings, we challenge the library community to make major changes to the Library of Congress Subject Headings system that have the potential to increase end-user understanding of subject headings.

According to Cutter (1904), the most important subject cataloging principle is consideration of the best interest of the catalog user. He stated (6): “The convenience of the public is always to be set before the ease of the cataloger.”

In the 90 years since Cutter laid down this rule, the Library of Congress Subject Headings (LCSH), the primary tool librarians consult for subject cataloging, has grown from a single volume listing a few thousand subject headings to a 5-volume set listing about 200,000 subject headings. In libraries throughout the country, librarians have produced tens of millions of unique cataloging records.
bearing subject headings drawn from this subject-cataloging tool. Yet not once in that time have catalogers asked library users whether they understood the subject headings assigned to cataloging records—nor did catalogers ask library users to suggest subject headings to represent the subject matter of the topics they seek.

In this article, we describe our large-scale empirical study of end-user understanding of subject headings. In the study, we focused on subdivided subject headings because the vast majority of subject headings in bibliographic files are subdivided (Drabenstott and Vizine-Goetz 1994). We addressed five research questions:

1. To what extent do end users understand subject headings?
2. Does end-user understanding vary based on subject heading context?
3. Does end-user understanding vary based on subject heading form?
4. Are there differences in levels of understanding between two types of end users (children and adults) and in levels of understanding for the different forms or contexts of subject headings?
5. What changes should be made to LCSH specifically and controlled vocabularies generally to improve end-user understanding of subject headings?

**Literature Review**

**The User and Usage**

Haykin (1951) established the principle of the reader as a focus. He stated (7):

"The reader is the focus in all cataloging principles and practice. All other considerations, such as convenience and the desire to arrange entries in some logical order, are secondary to the basic rule that the heading, in wording and structure, should be that which the reader will seek in the catalog, if we know or can presume what the reader will look under.

Chan (1986, 18) acknowledged that the meaning of Haykin's principle was "self-evident, but how to make it operational is not. The problem is delineating the user." Cutter and Haykin took different approaches to naming subjects in the catalog. Cutter (1904) recommended that public use be the guiding principle. Haykin (1951, 8) recommended "common usage or, at any rate, the usage of the class of reader for whom the material on the subject within which the heading falls is intended." In contrast to Cutter's straightforward approach, Haykin's approach gave the cataloger the freedom of naming subjects in the catalog based on the audience addressed by the material itself.

**NEW SUBJECT HEADINGS IN LCSH**

The addition of new subject headings to LCSH is the responsibility of an editorial group composed of Library of Congress (LC) staff members from the Cataloging Policy and Support Office and interested internal observers. The editorial group reviews proposals for changes to existing headings, i.e., "additions to, alterations in, or deletions of existing headings, heading/subdivision combinations, cross-references, or free-floating subdivisions" (Chan 1995, 146). The group also considers new subject headings and "deliberates on terminology (wording), cross-references, notes, compatibility with descriptive headings (if applicable), and conformity to existing patterns and broad policies governing LCSH" (Chan 1995, 146).

Until very recently, proposals for new headings and changes to existing headings emanated exclusively from catalogers at LC. The Vocabulary Improvement Project (Cochrane 1983) and an initiative sponsored by the Subject Analysis Committee (SAC) Subcommittee on New Subject Headings were pilot projects that demonstrated to LC that librarians at institutions other than LC could propose see references and new subject headings using the same procedures that LC librarians followed. Today, LC encourages librarians to submit new subject headings and see references by following the guidelines in the Subject Cataloging Manual: Subject Headings (SCM:SH) (Library of Congress 1991, H180–203). Despite such encouragement, very few subject heading proposals emanate from outside LC (Cooperative Subject Cataloging Project 1991).
In naming new subjects, LC catalogers face a more difficult task than their predecessors because of the diversity of today's catalog users. The decision to establish a new subject heading must take into consideration the best interest of users and the usage of the class of reader for whom the material is intended, and must avoid the use of terminology offensive to segments of the public. "When Cutter proposed his principle of the best interest of the user, he did not have a problem knowing users and usage because library users were a homogeneous group (Miksa 1983). Although today's user population is much more diverse than the user population was in Cutter's day, today's catalogers have tools to aid in the naming of subjects that their predecessors could not have imagined. For example, catalogers can examine an online catalog's transaction log to identify user queries that fail to retrieve records; they can then determine whether these queries should be represented in the controlled vocabulary as established headings or See references. Or catalogers can analyze catalog users' answers to online questionnaires in which users are asked questions about their interests, overall objectives, search requests, and the usefulness of search results. Before online systems, researchers and library practitioners did not have an accurate and systematic method of determining the subjects that users had difficulty finding in library catalogs.

END-USER UNDERSTANDING OF CATALOG INFORMATION

Bates (1977) demonstrated that users knowledgeable in a particular subject are as successful retrieving citations from the catalog as users without such knowledge. Interestingly, Bates found that the most successful users are those without subject expertise but with knowledge of the structure and content of the catalog.

Missing from the published literature on catalog use are studies of end-user understanding of subject headings. Researchers who examine the subject queries that online catalog users enter into catalogs provide us with an estimation of end-user understanding of catalog information. These researchers have demonstrated that users were not very successful at matching their queries for topical subjects or geographic names with the catalog's controlled vocabulary (Drabenstott and Vizine-Goetz 1994; Carlyle 1989), and even less successful at matching subject queries for personal names and combinations of topical subjects and names (Drabenstott and Vizine-Goetz 1994; Lester 1989).

Lilly (1954) provides some insight into end-user understanding of catalog information. He supplied students with the titles and authors of six books and instructed them to write down the subject headings under which they would expect to find each book. The percentages of correct student responses ranged from 2% to 64%.

IMPETUS FOR RESEARCH ON UNDERSTANDING SUBJECT HEADINGS

Following the publication of the LC Subject Subdivisions Conference's recommendations (Conway 1992), SAC established and charged the Subcommittee on the Order of LCSH Subdivisions to respond to the first of six recommendations of the LC Subject Subdivisions Conference. In this recommendation, it was suggested that the order of subject subdivisions be standardized for the purpose of simplifying subject cataloging: "If the cataloger chooses to apply subdivisions, the subdivisions should always appear in the following order: topical, geographic, chronological, form" (Conway 1992, 6). For three years, beginning with ALA's annual meeting in summer 1993, the subcommittee engaged in a multi-faceted study of the LCSH subject subdivisions system to ensure an informed decision regarding the future of subject subdivisions. Franz et al. (1994) reported on a pilot test of end-user understanding of subdivided subject headings that was conducted in connection with the subcommittee's work.

Pilot test researchers concluded that between 32% and 40% of end-user responses were correct meanings of subject headings. There was little difference between meanings for subject headings in the original and recommended orders of
subdivisions. However, users were more likely to ascribe a correct meaning to subject headings bearing few subdivisions (less than three) and few words (less than five). Although the findings of the pilot study were interesting, the study had limitations. First, it was difficult for the researchers to determine the meanings of subject headings because the catalogers who assigned meanings to the subject headings used in the study did not agree with one another and the researchers did not agree with the catalogers. Second, the generalizability of pilot study findings was suspect because researchers could not select a random sample of library users for inclusion in the study due to time constraints.

We designed the large-scale study described here to overcome the limitations of the pilot test. In this study, we enlisted an expert cataloger, with many years of experience in LC subject heading assignment, to determine the meaning of subject headings. We recruited a large number of respondents from public libraries to ensure that the study would not be plagued by generalizability questions. We adopted questions, format, procedures, and instructions from the pilot study and revised them based on pilot-study experience.

**DATA COLLECTION PROCEDURES**

We distributed self-administered questionnaires to children and adults at three public libraries in southeastern lower Michigan. The respondents provided demographic information in the questionnaires, and they were asked to write down the meaning of eight subject headings and to rate the certainty of each meaning.

**CONSTRUCTING SELF-ADMINISTERED QUESTIONNAIRES**

We selected a total of 24 LC subject headings for inclusion in the study from lists of frequently occurring and randomly selected subject headings from the OCLC Online Computer Library Center, Inc. Online Union Catalog (table 1). Team members deliberately chose subject headings. However, users were more likely to ascribe a correct meaning to subject headings bearing few subdivisions (less than three) and few words (less than five). Although the findings of the pilot study were interesting, the study had limitations. First, it was difficult for the researchers to determine the meanings of subject headings because the catalogers who assigned meanings to the subject headings used in the study did not agree with one another and the researchers did not agree with the catalogers. Second, the generalizability of pilot study findings was suspect because researchers could not select a random sample of library users for inclusion in the study due to time constraints.

We designed the large-scale study described here to overcome the limitations of the pilot test. In this study, we enlisted an expert cataloger, with many years of experience in LC subject heading assignment, to determine the meaning of subject headings. We recruited a large number of respondents from public libraries to ensure that the study would not be plagued by generalizability questions. We adopted questions, format, procedures, and instructions from the pilot study and revised them based on pilot-study experience.

**DATA COLLECTION PROCEDURES**

We distributed self-administered questionnaires to children and adults at three public libraries in southeastern lower Michigan. The respondents provided demographic information in the questionnaires, and they were asked to write down the meaning of eight subject headings and to rate the certainty of each meaning.

**CONSTRUCTING SELF-ADMINISTERED QUESTIONNAIRES**

We selected a total of 24 LC subject headings for inclusion in the study from lists of frequently occurring and randomly selected subject headings from the OCLC Online Computer Library Center, Inc. Online Union Catalog (table 1). Team members deliberately chose subject headings. However, users were more likely to ascribe a correct meaning to subject headings bearing few subdivisions (less than three) and few words (less than five). Although the findings of the pilot study were interesting, the study had limitations. First, it was difficult for the researchers to determine the meanings of subject headings because the catalogers who assigned meanings to the subject headings used in the study did not agree with one another and the researchers did not agree with the catalogers. Second, the generalizability of pilot study findings was suspect because researchers could not select a random sample of library users for inclusion in the study due to time constraints.

We designed the large-scale study described here to overcome the limitations of the pilot test. In this study, we enlisted an expert cataloger, with many years of experience in LC subject heading assignment, to determine the meaning of subject headings. We recruited a large number of respondents from public libraries to ensure that the study would not be plagued by generalizability questions. We adopted questions, format, procedures, and instructions from the pilot study and revised them based on pilot-study experience.

**DATA COLLECTION PROCEDURES**

We distributed self-administered questionnaires to children and adults at three public libraries in southeastern lower Michigan. The respondents provided demographic information in the questionnaires, and they were asked to write down the meaning of eight subject headings and to rate the certainty of each meaning.

**CONSTRUCTING SELF-ADMINISTERED QUESTIONNAIRES**

We selected a total of 24 LC subject headings for inclusion in the study from lists of frequently occurring and randomly selected subject headings from the OCLC Online Computer Library Center, Inc. Online Union Catalog (table 1). Team members deliberately chose subject headings. However, users were more likely to ascribe a correct meaning to subject headings bearing few subdivisions (less than three) and few words (less than five). Although the findings of the pilot study were interesting, the study had limitations. First, it was difficult for the researchers to determine the meanings of subject headings because the catalogers who assigned meanings to the subject headings used in the study did not agree with one another and the researchers did not agree with the catalogers. Second, the generalizability of pilot study findings was suspect because researchers could not select a random sample of library users for inclusion in the study due to time constraints.

We designed the large-scale study described here to overcome the limitations of the pilot test. In this study, we enlisted an expert cataloger, with many years of experience in LC subject heading assignment, to determine the meaning of subject headings. We recruited a large number of respondents from public libraries to ensure that the study would not be plagued by generalizability questions. We adopted questions, format, procedures, and instructions from the pilot study and revised them based on pilot-study experience.

**DATA COLLECTION PROCEDURES**

We distributed self-administered questionnaires to children and adults at three public libraries in southeastern lower Michigan. The respondents provided demographic information in the questionnaires, and they were asked to write down the meaning of eight subject headings and to rate the certainty of each meaning.

**CONSTRUCTING SELF-ADMINISTERED QUESTIONNAIRES**

We selected a total of 24 LC subject headings for inclusion in the study from lists of frequently occurring and randomly selected subject headings from the OCLC Online Computer Library Center, Inc. Online Union Catalog (table 1). Team members deliberately chose subject headings. However, users were more likely to ascribe a correct meaning to subject headings bearing few subdivisions (less than three) and few words (less than five). Although the findings of the pilot study were interesting, the study had limitations. First, it was difficult for the researchers to determine the meanings of subject headings because the catalogers who assigned meanings to the subject headings used in the study did not agree with one another and the researchers did not agree with the catalogers. Second, the generalizability of pilot study findings was suspect because researchers could not select a random sample of library users for inclusion in the study due to time constraints.

We designed the large-scale study described here to overcome the limitations of the pilot test. In this study, we enlisted an expert cataloger, with many years of experience in LC subject heading assignment, to determine the meaning of subject headings. We recruited a large number of respondents from public libraries to ensure that the study would not be plagued by generalizability questions. We adopted questions, format, procedures, and instructions from the pilot study and revised them based on pilot-study experience.

**TABLE 1**

<table>
<thead>
<tr>
<th>Subject Heading Set #1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Basketball—United States—Records</td>
</tr>
<tr>
<td>3. Locomotives—Germany—History</td>
</tr>
<tr>
<td>4. Music—500-1400—Philosophy and aesthetics</td>
</tr>
<tr>
<td>5. Indians of North America—New Mexico—Food</td>
</tr>
<tr>
<td>6. Spanish drama—18th century—History and criticism</td>
</tr>
<tr>
<td>7. Education—United States—Finance</td>
</tr>
<tr>
<td>8. Art, Modern—California—Los Angeles—20th century—Exhibitions</td>
</tr>
</tbody>
</table>

**Subject Heading Set #2**

<table>
<thead>
<tr>
<th>Subject Heading Set #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>9. Housing—United States—Law and legislation</td>
</tr>
<tr>
<td>10. Handicapped—Washington (State)—Seattle Metropolitan Area—Transportation</td>
</tr>
<tr>
<td>12. Organ music—17th century—Interpretation (phrasing, dynamics, etc.)</td>
</tr>
<tr>
<td>13. World War, 1939-1945—Regimental histories—Japan</td>
</tr>
<tr>
<td>14. English poetry—Old English, ca. 450-1100—Modernized versions</td>
</tr>
<tr>
<td>15. Music—Washington (D.C.)—History and criticism</td>
</tr>
<tr>
<td>16. Art, Modern—20th century—Germany—Berlin—Exhibitions</td>
</tr>
</tbody>
</table>

**Subject Heading Set #3**

<table>
<thead>
<tr>
<th>Subject Heading Set #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>17. Cattle—United States—Marketing</td>
</tr>
<tr>
<td>18. Combined sewers—Illinois—Chicago Metropolitan Area—Overflows</td>
</tr>
<tr>
<td>19. Art, Modern—20th century—Public opinion</td>
</tr>
<tr>
<td>20. Music—Africa—History and criticism—Bibliography</td>
</tr>
<tr>
<td>22. Music—Louisiana—New Orleans—History and criticism</td>
</tr>
<tr>
<td>23. Education—California—Finance</td>
</tr>
<tr>
<td>24. English poetry—Middle English, 1100-1500—Criticism, Textual—Congresses</td>
</tr>
</tbody>
</table>
headings that were likely to change in meaning when their subdivisions were reordered according to the recommendation of the LC Subject Subdivisions Conference to determine whether end users would in fact notice changes in meaning.

We constructed three separate sets of questionnaires corresponding to three sets of eight subject headings (i.e., subject headings 1–8, 9–16, and 17–24). Six different questionnaires made up each set. Questionnaires within sets varied in terms of the context in which subject headings were presented (i.e., alone, in bibliographic records, or in alphabetical browsing lists), and in terms of the order of subdivisions (i.e., original or recommended order) to minimize the order effect in data collection. Following each of the eight subject headings on questionnaires was a request for respondents to rate the certainty of the meaning they assigned to the subject heading on a scale from 1 (not at all certain) to 7 (very certain).

Recruiting Children and Adults

We recruited children and adults from three public libraries in southeastern lower Michigan—Flint Public Library, Bacon Memorial District Library, and Livonia Public Library. Flint Public Library has a professional staff of 32 and a collection of more than 500,000 items including government documents, video and audio tapes, microfilms, newspapers, and magazines. The library serves an immediate population of 139,000 within the city of Flint. Because it is the largest library across three counties, patrons come from all over mid-Michigan and beyond.

Bacon Memorial District Library serves 31,000 people in Wyandotte, south of Detroit on the Detroit River in the Downriver area. Downriver is made up of eleven old, established working-class towns, each with its own small library. Most of the population works in manufacturing, mainly in the automobile and steel industries. Wyandotte is a very stable community where often several family generations live close together. The library has a professional staff of 5 and a collection of about 70,000 items, including magazines, audio and video tapes, reference works on CD-ROM, and maps.

The Livonia Public Library has three branches and a reading room to serve Livonia, the eighth-largest city in Michigan with a population of 100,850. The library has a professional staff of 21 and a collection of almost 250,000 items. In addition to two public school districts, Livonia is home to Schoolcraft College and Madonna University.

Data collection procedures were similar from library to library. Interviewers stood at the main entrance of the library, introduced themselves to patrons who entered the library, and asked them to take part in the study. They told patrons the name and purpose of the study, explained the voluntary nature of participation, and told them that their complete participation would take ten to fifteen minutes. If patrons declined, interviewers thanked them for stopping, and the patrons continued on their way. Interviewers supplied participating patrons with an unmarked questionnaire, pencil, and eraser, and seated them at a nearby table, where the patrons completed the questionnaires. When finished, participants placed the completed questionnaires in a box provided for that purpose.

Recruiting children was not always as straightforward a process as recruiting adults. If interviewers were unsure whether library patrons were eighteen years old, the patrons were asked how old they were and the patrons were told that the same questionnaires were being given to both to adults and children but that interviewers needed to keep track of the number of each. If women entered the library with small children in tow, interviewers did not approach them to take part in the study because accompanying children could get bored, distracted, or annoyed, and cause their mothers to leave questionnaires incomplete.

Children often needed help with questionnaires. If children asked interviewers what a word meant, the interviewers would give them a simple definition. Interviewers found that it was impossible for children under age ten to complete questionnaires. When inter-
viewers did give questionnaires to children ten years old or younger, the children usually returned them and said, "I can't do this" or "This is too hard." Some children asked their older sisters, brothers, or parents to help them read the words. Parents read hard words and, in rare cases, provided definitions for words in subject headings. When children asked interviewers for clarification on instructions, interviewers told them to try to put sure that two experts with similar experiences to the first, second, and third sets of responses.

Determining how to arrive at correct meanings for the subdivided subheadings used in the study was difficult because we knew that these meanings would be used to judge all the meanings provided by children and adults. We considered using a consensus of responses from professional librarians to determine the meaning of subdivided subject headings. However, results from the pilot test of subject heading understanding demonstrated that professional librarians—both reference and technical services librarians—did not agree on the meaning of subdivided subject headings (Franz et al. 1994).

After considerable deliberation, we enlisted a single subject-cataloging expert with many years of experience in LC subject heading assignment. We consulted a subject-cataloging expert at the University of Michigan who had more than twenty-five years of experience in LC subject heading practice and many years of service on professional committees in the area of subject access. Not only was the expert familiar with our objectives, she had read the proposal to the organization that funded the study and knew that she would be reviewing subject headings that were not "correct" in terms of the order of subdivisions. She completed all 18 versions of questionnaires in several settings and did not compare subject headings between questionnaires. Although she did not know which headings were correct or incorrect, on occasion she was able to guess which headings were incorrect based on her knowledge of LC subject heading practice.

We undertook a reliability study to ensure that two experts with similar experience would agree on correct meanings. In every case, the second expert's responses corresponded closely in meaning to the responses of the first expert. There were differences in language and syntax but, overall, the second expert gave responses that would be considered correct responses.

**ANALYZING COLLECTED DATA**

**CATEGORIZING USER-ASSIGNED MEANINGS**

Three research team members examined user-assigned meanings from completed questionnaires and placed them into correct and incorrect meaning categories and subcategories. Coders first read the expert-supplied meaning and paid close attention to its syntax, language, and meaning. They then read the user-assigned meaning on the completed questionnaire and compared it to the expert-supplied meaning. They looked for similarities and differences in language or word choice, syntax, and meaning. They determined whether meanings were "correct" or "incorrect." They then assigned one to two codes to indicate the nature of "correctness" or "incorrectness" of the user-assigned meaning. Details on "correct" and "incorrect" subcategories follow.

**CORRECT SUBCATEGORIES**

There were two correct subcategories that coders assigned to correct meanings only: (1) used same language as expert's meaning, and (2) used language that was different from expert's language.

*Same Language.* If the comparison between the expert-supplied meaning and the user-assigned meaning revealed no
differences in word choice, syntax, or meaning, coders assigned the user’s meaning to the “Same Language” subcategory. An example was the expert-supplied meaning “history of locomotives in Germany” and user-assigned meaning “a history of locomotives in Germany” for the subdivided, reordered subject heading “Locomotives—History—Germany.” The coder assigned the “Same Language” subcategory because the user’s meaning matched the language of the expert’s meaning letter-for-letter except for the initial article.

**Different Language.** If coders determined that the comparison between the expert-supplied meaning and the user-assigned meaning revealed that the respondent used different language to capture the same meaning as the expert-supplied meaning, coders assigned the “Different Language” subcategory. An example was the user-assigned meaning “records (statistics) for U.S. basketball” for the subject heading “Basketball—United States—Records.” The expert-supplied meaning for this subject heading was “records of U.S. basketball.”

**Correct or Incorrect Subcategories**

There were four subcategories that coders assigned to correct or incorrect meanings: (1) read in one concept, (2) read in more than one concept, (3) used different syntax, and (4) combinations of two correct or two incorrect subcategories. Sometimes the read-in concepts did not affect the meaning of respondent-assigned meanings to the extent that they were incorrect compared to expert-supplied meanings and at other times the read-in concepts resulted in incorrect meanings.

**Read in One Concept.** Coders assigned the subcategory “Read in One Concept” when their comparison of the user-assigned and expert-supplied meanings revealed that the end user added a concept (i.e., a word or phrase). Coders qualified this subcategory to indicate whether the user’s addition of a concept resulted in a correct or incorrect meaning for the subdivided subject heading. An example of a user’s correct meaning in this category was “food used (eaten, cooked, etc.) of Indians in New Mexico” for the subject heading “Indians of North America—New Mexico—Food.” The expert-supplied meaning was “food of the Indians of New Mexico.” The user’s addition of more than one concept—“used,” “eaten,” and “cooked”—made this meaning appropriate for assignment to the “Read in More Than One Concept (Correct)” subcategory.

An incorrect example in this subcategory follows. The expert-supplied meaning for the reordered subject heading “Cattle—Marketing—United States” was “marketing of cattle in the U.S.” An end user gave the meaning “current or historical cases and techniques for marketing cattle and or beef financial aspects of U.S. education” for the subject heading in original order “Education—United States—Finance.” The expert-supplied meaning was “finance of U.S. education.” The coder assigned the code “Read in One Concept (Correct)” because the user’s meaning matched the expert-supplied meaning even though the former contained the concept “financial aspects” that was different from “finance” in the latter.

Here is an example of a meaning that the coder determined was incorrect for the “Read in One Concept” subcategory. The expert-supplied meaning was “transportation of the handicapped in the Seattle (Washington) metropolitan area” for the subject heading “Handicapped—Washington (State)—Seattle Metropolitan Area—Transportation.” The coder assigned the “Read in One Concept (Incorrect)” subcategory to the end user’s meaning because the user added the concept “public transportation.”

**Read in More Than One Concept.** Coders assigned the subcategory “Read in More Than One Concept” when their comparison of the user-assigned and expert-supplied meanings revealed that the user added more than one concept (i.e., words or phrases). Coders qualified this subcategory to indicate whether the user’s addition of concepts resulted in a correct or incorrect meaning for the subdivided subject heading. An example of a user’s correct meaning in this category was “food used (eaten, cooked, etc.) of Indians in New Mexico” for the subject heading “Indians of North America—New Mexico—Food.” The expert-supplied meaning was “food of the Indians of New Mexico.” The user’s addition of more than one concept—“used,” “eaten,” and “cooked”—made this meaning appropriate for assignment to the “Read in More Than One Concept (Correct)” subcategory.
products” to which the coder assigned “Read in More Than One Concept (Incorrect)” because the user’s meaning was semantically different from the expert’s meaning and included more than one concept (i.e., “current or historical cases,” “techniques,” and “beef products”).

**Different Syntax.** Coders assigned the “Different Syntax” subcategory when their comparison of the expert-supplied meaning and the respondent-assigned meaning revealed that the respondent used the same language but different syntax to capture the same meaning as the expert-supplied meaning. Coders qualified this subcategory to indicate whether the syntax made the meaning correct or incorrect. An example of a correct meaning with different syntax was the user-assigned meaning “Washington (D.C.) music—history and criticism” for the subject heading with subdivisions in recommended order “Music—History and criticism—Washington (D.C.).” The expert-supplied meaning for this subject heading was “history and criticism of Washington (D.C.) music.”

Most of the time, different syntax meant an incorrect meaning. An example was the user-assigned meaning “history in Germany of locomotives” given to the subdivided subject heading in original order “Locomotives—Germany—History.” The expert-supplied meaning was “history of locomotives in Germany.” In this case, the difference in syntax changed the meaning.

**Combinations.** At times, coders found it impossible to assign only one subcategory to correct or incorrect meanings because more than one situation occurred. For example, the coder assigned “Left Out One Concept (Incorrect)” and “Read in More Than One Concept (Incorrect)” to the user-assigned meaning “a history of Jewish immigrants in Detroit where the majority came from church, work, social problems, and how they are resolving them” for the subject heading “Jews—Michigan—Detroit—History—20th century.” In this case, the user left out the “20th century” element and read in concepts such as “immigrants,” “church,” and “social problems.”

**Incorrect Subcategories**

There were three subcategories that coders assigned to incorrect meanings only: (1) left out one concept, (2) left out more than one concept, (3) no response.

**Left Out One Concept.** Coders assigned the “Left Out One Concept” subcategory when their comparison of the user-assigned and expert-supplied meanings revealed that the user had omitted a concept. Meanings assigned to this subcategory were always incorrect because the omission resulted in an incorrect meaning. For example, the expert-supplied meaning for the subject heading “Housing—United States—Law and legislation” was “law and legislation of housing in the U.S.” Coders assigned the “Left Out One Concept” subcategory to the user-assigned meaning “laws on housing in the U.S.” because the “legislation” concept was omitted and the omission resulted in an incorrect meaning.

**Left Out More Than One Concept.** Coders assigned the “Left Out More Than One Concept” subcategory when their comparison of the user-assigned and expert-supplied meanings revealed that the user had omitted more than one concept. The omission of concepts always changed the meaning of subject headings, thus, this was a subcategory for incorrect meanings. For example, the expert gave the meaning “exhibitions of 20th century Los Angeles (California) modern art” to the subject heading “Art, Modern—California—Los Angeles—20th century—Exhibitions.” Here are several respondent-assigned meanings missing more than one concept:
- California 20th century
- Art different places
- The new art
- Art in California in the 20th century

**No Response.** When respondents failed to assign meanings to questionnaires and left the response blank, the coder considered this a null response and coded it as “No Response.” Sometimes, users wrote messages telling us about their difficulties. For example, the expert-supplied meaning for the reordered subject heading “English Poetry—Modernized Versions—Old English, ca.
450-1100" was "modernized versions of old English (ca. 450-1100) poetry." The respondent wrote down "basically nothing, since I don’t know what ‘modernized versions’ means." Such a meaning was only appropriate for the "No Response" category because the respondent told us why he or she could not supply a meaning.

**Characteristics of Participating Children and Adults**

We met our goal of recruiting 48 children and 48 adults at each of the three participating libraries. A total of 144 children and 144 adults took part in the study. Overall, about two-thirds of library patrons who completed questionnaires were female and one-third were male. At Livonia, almost 80% of participating library patrons were female. The largest percentage of participating males came from Wyandotte, where 41% of respondents were male.

Table 2 shows ages reported by adult library patrons. (The total number of adults does not amount to 144 in table 2 because some respondents failed to answer the question about age.)

At Flint and Wyandotte, more than half of adults were 18 to 40 years old. At Livonia, only 26% were in this age range. More than half of Livonia's library patrons were over 50. At Flint and Wyandotte, only 18% and 13% of adults were in this age range, respectively.

Table 3 shows children's ages. (The total number of children does not amount to 144 in table 3 because some respondents failed to answer the question about their age.)

The largest percentages of young children came from Flint, where a little more than a third of children were 12 years old or
younger. At Wyandotte and Livonia, 17% and 21%, respectively, of children were 12 years old or younger. About half of Wyandotte children were 16 or 17 years old. Overall, all but one age category (14 years old) registered double-digit percentages.

Overall, one-third of adult library patrons reported that they had attended but not graduated from college. A little over half (52%) of participating adults were college graduates. Small percentages (1% and 14%) of adults had completed only junior and senior high school, respectively. The largest percentage (67%) of adult library patrons who had a college degree came from Livonia.

Only 5% of children in the study had graduated from high school or had had some college. A large percentage of children (38%) had completed elementary school and a larger percentage of children (57%) had completed junior high school.

Questionnaires allowed library patrons to write down a word or phrase that described their profession. About 16% of adults failed to write down such a word or phrase. We consolidated professions into 12 broad categories. Categories that described 4% or more of the adults participating in the study were:
- Students (43%)
- Retired (7%)
- Education (7%)
- Homemakers (5%)
- Science, technology, and computer fields (4%)
- Tradespersons, e.g., autoworkers, electricians, maintenance workers, cooks (4%)

At all three libraries, patrons visited the library on a weekly or monthly basis. A breakdown of these percentages for children and for adults shows the same pattern of weekly or monthly library use.

We intended our analysis of demographic information about study participants to be descriptive of the people who took part in the study. Both we and the administrators of participating libraries were sensitive about analyses that compared particular subpopulations of a library's clientele and subject heading understanding because we could not promise that changes to the subject subdivision system could be made that would increase understanding by a particular subpopulation. Consequently, we limited the data analyses in this study to a comparison of subject heading understanding between children and adults generally, and did not extend to other factors such as respondents' ethnic background, gender, socio-economic status, etc.

A STATISTICAL ANALYSIS OF USERS' MEANINGS

CORRECT AND INCORRECT MEANINGS

Figure 1 shows overall percentages of correct and incorrect meanings for children and adults. Percentages of correct meanings were quite different for the two respondent types. Children provided correct meanings for 31% of the test headings, while adults provided correct meanings for 39% of the headings.

To compare the performance of children and adults in terms of assigning correct meanings to subject headings, we submitted collected data to a four-way analysis of variance (ANOVA) with Library, Type of Respondent, and Context as between-subject factors and with Subdivision Order as a within-subject factor. Table 4 summarizes the result of the analysis for main effects. The upper limit for mean correct meanings was 4 because individual respondents gave meanings to four subject headings in original order and to
TABLE 4
RESULTS OF FOUR-WAY ANOVA FOR CORRECT MEANINGS

<table>
<thead>
<tr>
<th>Results for Type of Respondent</th>
<th>Mean = 1.24</th>
<th>Standard deviation = 1.18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children</td>
<td>Mean = 1.57</td>
<td>Standard deviation = 1.13</td>
</tr>
<tr>
<td>Adults</td>
<td>H0: Type of Respondent effect — $F_{(1, 270)} = 10.36$ — Significance = .001*</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Results for Library</th>
<th>Mean = 1.69</th>
<th>Standard deviation = 1.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flint</td>
<td>Mean = 0.95</td>
<td>Standard deviation = 1.06</td>
</tr>
<tr>
<td>Wyandotte</td>
<td>Mean = 1.57</td>
<td>Standard deviation = 1.13</td>
</tr>
<tr>
<td>Livonia</td>
<td>H0: Library effect — $F_{(2, 270)} = 19.82$ — Significance = .000*</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Results for Context</th>
<th>Mean = 1.43</th>
<th>Standard deviation = 1.18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alone</td>
<td>Mean = 1.27</td>
<td>Standard deviation = 1.17</td>
</tr>
<tr>
<td>Bibliographic record</td>
<td>Mean = 1.51</td>
<td>Standard deviation = 1.12</td>
</tr>
<tr>
<td>Alphabetic list</td>
<td>H0: No Context effect — $F_{(2, 270)} = 1.90$ — Significance = .152</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Results for Subdivision Order</th>
<th>Mean = 1.45</th>
<th>Standard deviation = 1.19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original order</td>
<td>Mean = 1.36</td>
<td>Standard deviation = 1.13</td>
</tr>
<tr>
<td>Recommended order</td>
<td>H0: No Subdivision Order effect — $F_{(1, 270)} = 1.43$ — Significance = .234</td>
<td></td>
</tr>
</tbody>
</table>

*Significant at the .05 level.

There were significant results for two factors: Type of Respondent and Library. There were no other main effects or interactions that were significant at the .05 level. With respect to Type of Respondent, the means for children and adults varied by a third of a point. Adults therefore performed significantly better than children in terms of assigning correct meanings to subdivided subject headings.

With respect to Library, the means for assigning correct meanings for respondents at Flint and Livonia were about the same; however, the mean for Wyandotte respondents was about two-thirds of a point lower than the means for respondents at the other locations. Because respondents at the three participating libraries examined different sets of subject headings, it was impossible to attribute the effect to the different libraries or the different subject headings enumerated on questionnaires. Thus, no conclusions could be drawn about the significant effect because of confounding factors.

We found no significant effect for Context. Means for the three contexts were slightly different—respondents did best when they assigned meanings to subject headings in alphabetical browsing lists and they did worst when they assigned meanings to subject headings in bibliographic records, but there was a difference of hardly one-quarter of a point between the means. We concluded from this analysis that Context had no effect on respondents’ ability to assign correct meanings.

We found no significant effect for Subdivision Order. Less than a tenth of a point separated mean correct meanings for subject headings in the original order and for subject headings in the recommended order. Thus, children and adults performed about as well in terms of assigning correct meanings whether they examined subject headings in the original order or in the recommended order of subdivisions. This surprised us. The impetus for this research was a recommendation to standardize the
order of subdivisions. Some members of
the ALA subcommittee studying the prob-
lem of subdivision order felt that standard-
zizing the order of subdivisions might con-
fuse library users and adversely affect their
ability to give correct meanings to subdivi-
ded subject headings. The result here
demonstrated that changing the order of
subject subdivisions did not have a negative
impact on library users and their under-
standing of subject headings.

**A FAILURE ANALYSIS OF END USERS’ MEANINGS**

This failure analysis of end users’ mean-
ings features the specific reason or rea-
sons why end users’ meanings were cor-
rect or incorrect using the subcategories into which team members assigned users’
meanings.

**CORRECT MEANING SUBCATEGORIES**

Figure 2 shows that the patterns of percent-
ages of correct subcategories for children’s
and adults’ meanings were rather different.
Children’s correct meanings usually were
exact representations of the expert’s mean-
ings or used different syntax. Children used
different language but syntax and exact
matches were more likely. Children rarely
gave correct meanings that read in con-
cepts. In contrast, the largest percentage of
adults’ correct meanings used different lan-
guage. Adults also used different syntax and
matched the expert’s meanings exactly. On
occasion, adults gave correct meanings that
read in concepts.

Here are three examples of subject
headings for which both children and
adults matched the expert’s meanings.
• Locomotives—Germany—History / history of locomotives in Germany
• Cattle—United States—Marketing / marketing of cattle in the United
States
• Jews—Egypt—Politics and govern-
ment / politics and government of
Jews in Egypt

Meanings assigned to the “Different
Language (Correct)” subcategory ac-
counted for 21% of children’s correct
meanings and 35% of adults’ correct
meanings. Table 5 lists subject headings, ex-
ert-supplied meanings, and examples of
library users’ meanings in this subcategory.

Meanings assigned to the “Different
Syntax (Correct)” subcategory accounted
for 44% of children’s correct meanings
and 25% of adults’ correct meanings.
Table 6 gives examples of library users’
meanings in this subcategory.
Both children and adults seldom gave correct meanings that read in concepts. Examples for the subject heading “English poetry—Modernized versions—Old English, ca. 450–1100” to which the expert assigned the meaning “modernized versions of old English (ca. 450–1100) poetry” are:

- English poetry translated to modern version from Old English, ca. 450–1100
- Modern English versions of English poetry written between 450–1100
- English poetry between 450–1100 ca. that has been updated so that [it] is readily understandable to the average Joe

The first two meanings read in the “translated” and “written” elements and the last meaning read in the story about the “average Joe.” These read-in elements did not make the meanings incorrect but they did clarify the meaning.

**INCORRECT MEANING SUBCATEGORIES**

Figure 3 shows percentages of incorrect subcategories for library users’ meanings. In this instance, patterns for children and adults were not that much different. Percentages were highest for “Different Syntax” but there were also high percentages of incorrect meanings for leaving out one or more concepts and reading in one concept. There were small percentages for reading in more than one concept, combinations, and no responses.

---

**TABLE 5**

**CORRECT MEANINGS IN THE “DIFFERENT LANGUAGE” SUBCATEGORY**

<table>
<thead>
<tr>
<th>Subject Headings</th>
<th>Expert’s Meaning</th>
<th>Library Users’ Meanings</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1: Basketball—Records—United States</td>
<td>Records of U.S. basketball</td>
<td>United States basketball teams, scores, and other game stuff</td>
</tr>
<tr>
<td></td>
<td>Records of basketball in repositories in the U.S.</td>
<td>Basketball statistics in the U.S.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>History of Jews who lived in Detroit, Mich., during the 20th century</td>
</tr>
<tr>
<td>#17: Cattle—United States—Marketing</td>
<td>Marketing of cattle in the U.S.</td>
<td>The selling of cattle in the U.S.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Information on marketing of cattle in the U.S.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>How cattle are marketed in the U.S.</td>
</tr>
</tbody>
</table>

---

**TABLE 6**

**CORRECT MEANINGS IN THE “DIFFERENT SYNTAX” SUBCATEGORY**

<table>
<thead>
<tr>
<th>Subject Headings</th>
<th>Expert’s Meaning</th>
<th>Library Users’ Meanings</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1: Basketball—United States—Records</td>
<td>Records of U.S. basketball</td>
<td>U.S. basketball records</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The basketball records of the U.S.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Basketball records in the U.S.</td>
</tr>
<tr>
<td>#4: Music—Philosophy and aesthetics—500-1400</td>
<td>Philosophy and aesthetics of music for the time period 500-1400</td>
<td>Music philosophy and aesthetics from 500-1400</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The philosophy and aesthetics of music in the time period 500-1400</td>
</tr>
<tr>
<td>#7: Education—United States—Finance</td>
<td>Finance of education in the U.S.</td>
<td>U.S. finances education</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Education finance in the U.S.</td>
</tr>
</tbody>
</table>
The response pattern in figure 3 summarizes the percentages of incorrect meaning subcategories across all three sets of subject headings; however, response patterns were quite different for each of the three sets of subject headings. For example, children's meanings were characterized by a high percentage (43%) of “Left Out More Than One Concept” meanings for the second set of subject headings. Adults' meanings featured moderately high percentages of “Left Out One Concept” (22%), “Left Out More Than One Concept” (23%), and “Read in One Concept” (24%) meanings for the third set of subject headings. The sections that follow examine the specific problems children and adults experienced.

**Syntax Problems**

Incorrect meanings with “Different Syntax” were typical of subject headings to which the expert gave different meanings to the two different forms of subject headings. Users assigned incorrect meanings to subject headings that would have been judged correct for the order of the subject heading. For example, the expert supplied the meaning “public opinion of 20th century modern art” to the subject heading “Art, Modern—20th century—Public opinion” in the original order of subdivisions and she assigned the meaning “20th century public opinion of modern art” to the subject heading “Art, Modern—Public opinion—20th century” in the recommended order of subdivisions. Respondents assigned the former meaning to the latter subject heading and vice versa. Table 7 lists three more subject headings that gave users the same type of syntax problem.

This type of syntax problem did not occur for every subject heading for which changing the order of subdivisions also changed the subject heading's meaning. Nor did all the children or all the adults who gave meanings to these subject headings generate incorrect meanings with this particular type of syntax problem—but it did happen with some frequency.

“Different Syntax (Incorrect)” meanings sometimes introduced entirely new meanings for subject headings that were not amongst the one or more meanings supplied by the expert. For example, the expert gave the subject heading “Education—United States—Finance” in the original and recommended orders only one meaning—“finance of U.S. education.” Examples of meanings with syntax
problems that gave entirely new meanings to this subject heading were:
- U.S. education in finance
- Finance education in the U.S.
- U.S. finance education
- Education and finance in the U.S.

LEAVING OUT CONCEPTS
Incorrect meanings for "Leaving Out One Concept" were more typical of adults than children. Children also left out concepts but they were more likely to leave out two or more concepts. Table 8 features the many meanings assigned to four subject headings that were missing more than one concept. Most of the examples in table 8 came from children.

Users consistently missed the "combined" element in the "Combined sew-

ers" subject heading. They usually glossed over this concept by referring to "sewer problems" or "information on sewers" in their meanings. The "regimen-
tal histories" element was consistently missing from the examples for the "Ja-
pnish regimental histories" heading. Some respondents also left out the "Japanese" element.

Users did not consistently omit certain topics from their meanings for the other two subject headings. For example, meanings for the "Housing" heading left out one or more of the four concepts in this subject heading, i.e., "housing," "law," "legislation," and "U.S." Overall, we could characterize incorrect meanings for "left out concepts" in two ways: (1)
TABLE 8
INCORRECT MEANINGS IN THE "LEFT OUT MORE THAN ONE CONCEPT" SUBCATEGORY

<table>
<thead>
<tr>
<th>Subject Headings</th>
<th>Expert's Meanings</th>
<th>Library Users' Meanings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing—United States—Law and legislation</td>
<td>Law and legislation of U.S. housing</td>
<td>Types of housing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congress</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Housing in the U.S.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Government</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The laws for a house</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(rental or own)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Law and legislation</td>
</tr>
<tr>
<td>World War, 1939–1945—Regimental histories—Japan</td>
<td>Japanese regimental histories of World War (1939–1945)</td>
<td>WWII—concerning Japan</td>
</tr>
<tr>
<td></td>
<td>Regimental histories of World War (1939–1945) located in Japan</td>
<td>Boundaries war related</td>
</tr>
<tr>
<td></td>
<td></td>
<td>History of military</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Things that went on in WWII</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wars</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The war against Japan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Japan and the world wars</td>
</tr>
<tr>
<td>Music—Washington (D.C.)—History and criticism</td>
<td>History and criticism of Washington (D.C.) music</td>
<td>Music reflecting history</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Music peculiar to Washington, D.C.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congress</td>
</tr>
<tr>
<td></td>
<td></td>
<td>History of music</td>
</tr>
<tr>
<td>Combined sewers—Illinois—Chicago Metropolitan Area—Overflows</td>
<td>Overflows of combined sewers in the Chicago (Ill.) metropolitan area</td>
<td>Where there is a “problem” with overflow of sewers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Information on the sewers of Chicago, Ill.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sewer flow in Chicago</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The unique sewer problems of the Chicago area</td>
</tr>
</tbody>
</table>

meanings that captured two or more of the several specific concepts present in the subject, e.g., “music peculiar to Washington, D.C.” and “history of music” for the “Music” subject heading in table 8, and (2) meanings that described the heading in one or two broad terms, e.g., “the war against Japan,” “Japan and the world wars,” and “wars” for the “World War” subject heading in table 8.

READING IN CONCEPTS
Reading in concepts occurred infrequently. The subject headings to which some patrons assigned “read-in” meanings were also likely to inspire other patrons to generate meanings with syntax errors or to leave out one or more concepts.

The concepts that patrons read into their meanings were rarely concepts that were present in the bibliographic records or alphabetical browsing lists that were displayed along with the subject heading. An examination of user-assigned meanings for the subject heading “Basketball—United States—Records” resulted in examples of meanings that contained read-in concepts in addition to elements present in the expert’s meaning “Basketball records in the U.S.” Examples of users’ meanings were:
- The records of the basketball players in the U.S.
- Athletic record holders in basketball
- I would find facts relating to basketball records made or broken in the U.S.
- NBA player and team statistics
The read-in concepts in these user-assigned meanings probably reflected users' own personal knowledge of and experience with this subject. Some read-in concepts might have also been inspired by the bibliographic record's title (*Basketball Statistics: Top Players and Teams by Game, Season, and Career*) because the users mentioned basketball players, teams, and the NBA. Here are more user-assigned meanings with read-in concepts for the subject heading "Cattle—United States—Marketing." The expert's meaning for this subject heading was "marketing of cattle in the U.S."

- Making money—dealing in cattle/ U.S.
- How to market cattle profitably in U.S. by knowing cattle cycles
- Law and legislation regarding cattle marketing in U.S.
- How the USDA goes about marketing cattle

The title of the bibliographic record (*Cattle Cycles: How to Profit From Them*) could have inspired the first two meanings listed above. Patrons who formulated the last two titles above did not see the subject heading in a bibliographic record. They could have added the phrases "law and legislation" and "USDA" based on their own experiences and knowledge.

**No Response**

We categorized more than 10% of the responses for seven subject headings into the "no response" subcategory. Because we had about 50 meanings per subject heading from children and 50 meanings per subject heading from adults, this meant that five children or adults failed to provide meanings for these subject headings. These seven subject headings are listed in table 9 along with the percentages of "no responses" and the number of meanings for the six representations of the subject heading.

Most percentages were around 12%, but there were two percentages that accounted for much more. Two of the seven subject headings featured only one meaning while another two featured as many as three meanings.

We looked at the other incorrect meanings users assigned to the subject headings in table 9. Respondents were also likely to assign to these subject headings incorrect meanings that either left out or read in one or more concepts. Both children and adults typically left out one or more concepts when assigning meanings to the subject heading "Jews—Germany—Berlin—Intellectual life—Congresses." Examples of children's meanings were:

- Religion
- About different kinds of people
- A Jew's life
- About Germany
- How Jewish people live their life

Read-in concepts for these headings were typical of the incorrect meanings users gave to the subject heading "Educa-

---

### Table 9: Subject Headings and Blank Responses

<table>
<thead>
<tr>
<th>Subject Heading</th>
<th>No Response (%)</th>
<th>No. of Meanings</th>
</tr>
</thead>
<tbody>
<tr>
<td>#4: Music—Philosophy and aesthetics—500-1400</td>
<td>35</td>
<td>1</td>
</tr>
<tr>
<td>#24: English poetry—Middle English, 1100-1500—Criticism, Textual—Congresses</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>#21: Jews—Egypt—Politics and government</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>#11: Jews—Germany—Berlin—Intellectual life—Congresses</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>#23: Education—California—Finance</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>#15: Music—Washington (D.C.)—History and criticism</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>#22: Music—Louisiana—New Orleans—History and criticism</td>
<td>10</td>
<td>3</td>
</tr>
</tbody>
</table>
tion—California—Finance.” Some of the concepts children and adults read into their meanings might have come from the bibliographic record in which they saw this subject heading which contained the concepts “school” and “budget.” Examples of incorrect meanings with read-in concepts were:

- Cost of education in California
- California school budget
- How much they waste money on education in California
- How to finance your education in California
- Education on young kids and teens (in California)

**CERTAINTY OF CORRECT MEANINGS**

Respondents were not only asked to write down the meaning of subject headings, they were asked to rate the certainty of their meanings on a scale from 1 (not at all certain) to 7 (very certain). Overall, children gave certainty scores that averaged 5.05 for correct meanings. They gave certainty scores that averaged 4.15 for incorrect meanings. Their average certainty score for correct meanings was higher than their average certainty score for incorrect meanings and a little less than one point separated the two scores. Figure 4 shows average certainty scores that children gave to correct and incorrect meanings of subject headings for each of the three sets of subject headings. Children’s certainty scores for correct meanings were always higher than their certainty scores for incorrect meanings. The difference between certainty scores for correct and incorrect meanings varied by as little as about a half point (set 3) and as much as about one point (sets 2 and 3).

**Certainty scores for the two orders of subdivisions or three contexts of subject headings mirrored the general trend that certainty scores for correct meanings were greater than such scores for incorrect meanings. This was not true across the board for certain contexts and orders of subdivisions, as percentages of correct and incorrect meanings flip-flopped.**

Overall, adults gave certainty scores that averaged 5.70 for correct meanings.

**Figure 4. Children’s Certainty Scores.**

They gave certainty scores that averaged 5.08 for incorrect meanings. The average certainty scores that adults gave to correct and incorrect meanings were three-quarters of a point to almost a full point higher than the average certainty scores that children gave to correct and incorrect meanings. Figure 5 shows average certainty scores that adults gave to correct and incorrect meanings of subject headings for each of the three sets of subject headings.

**Adults’ Certainty Scores for Correct Meanings were higher than their certainty scores for incorrect meanings for the three subject headings sets. The difference between certainty scores for correct and incorrect meanings varied by as little as about a half point (sets 1 and 3) and as much as about one point (set 2).**

**Figure 5. Adults’ Certainty Scores.**
Certainty scores for the two orders of subdivisions or three contexts of subject headings again mirrored the general trend that certainty scores for correct meanings were greater than such scores for incorrect meanings. Sometimes the difference between two scores was very low (just a few hundredths of a point separating the two certainty scores) and sometimes the difference exceeded a point.

Generally unusually high or low certainty scores were not associated with a particular correct or incorrect meaning code. We found one exception. Unusually low certainty scores were almost always associated with both incorrect left-out subcategories.

**Discussion and Implications of Study Findings**

In this section, we discuss the findings in terms of the five research questions we sought to answer.

1. **Do users understand subject headings?** Overall, about 36% of the meanings end users gave to subdivided subject headings were correct. Some readers might conclude that this percentage was low. But the subdivided subject headings to which end users assigned meanings were complex. Some subject headings had as many as three to four subdivisions. Other subject headings featured one or more main headings or subdivisions that were inverted, contained more than one phrase, or featured subdivisions with qualifiers.

   In addition, the rules that coders followed to determine whether meanings were correct and incorrect were rather stringent. Study participants had to provide meanings that contained all elements present in the expert’s meaning. The syntax of their meanings had to match the syntax of the meanings supplied by the expert. If users formulated meanings that included additional elements besides those present in expert-supplied meanings, their meanings were usually marked incorrect.

   Furthermore, we have no basis for comparison. To our knowledge, this is the first large-scale study of end-user understanding of subject headings. If findings from other studies were available, we might be able to compare percentages of correct meanings and determine just how low or how high these scores were.

2. **Does end-user understanding vary based on subject heading context?** Mean correct meanings for the three contexts varied very little. End users did best (mean = 1.51) when they assigned meanings to subject headings in alphabetical browsing lists, and they did worst (mean = 1.27) when they assigned meanings to subject headings in bibliographic records (table 4). But there was hardly one-quarter of a point difference between the means, and the difference was not significant. On occasion, we noticed that library patrons added concepts to their meanings that occurred in bibliographic records, but this did not happen very frequently. When it did happen, adults rather than children were usually the ones giving meanings with read-in concepts.

3. **Does end-user understanding vary based on subject heading order?** This research question was one of our most important because the impetus for this study was a recommendation to standardize the order of subdivisions (Conway 1992). Before implementing the recommendation, librarians wanted to determine whether reordering subdivisions would have a negative effect on end users’ understanding of subdivided subject headings. Some librarians expected that end users would have more problems understanding the meaning of subject headings in the recommended order than in the original order.

   Less than a tenth of a point separated the two mean correct meanings for subject headings in the original (1.45) and recommended (1.36) orders. Analysis of Variance demonstrated that library users performed about as well in terms of assigning correct meanings whether they examined subject headings in original order or in the recommended order of subdivisions (table 4).

4. **Are there differences in understanding between children and adults?** Figure 1 showed that percentages of correct meanings were quite different for children and adults—31% of children’s
meanings were correct and 39% of adults' meanings were correct. Although mean correct meanings for children (1.24) and adults (1.57) varied by only a third of a point, Analysis of Variance demonstrated that there was a significant difference between the two means (table 4). Adults did significantly better than children in terms of assigned correct meanings to subdivided subject headings.

Findings about certainty scores were heartening. They demonstrated that both children and adults had less confidence in their incorrect meanings than in their correct meanings.

5. What changes should be made to LCSH and other controlled vocabularies to improve user understanding? Do the data and analyses presented here suggest making major changes to the existing system of subject headings based on end users' lack of understanding and difficulty with subject headings? We feel that the library community needs to grapple with the issues involved with answering this question and make some important decisions.

The statistical analysis did not result in a significant main effect for Subdivision Order. Thus, children and adults performed about as well in terms of assigning correct meanings regardless of the order of subdivisions in subject heading strings. Because subdivision order did not negatively affect end-user understanding, we recommend standardizing the order of subject subdivisions. Standardizing subdivision order would simplify cataloging and save money. Library school faculty and technical services staff would not have to spend time training students and staff how to order the subdivisions in subject heading strings. Cataloging would be streamlined because staff would no longer spend time determining the order of subject subdivisions. Instead, they would build strings based on a standardized order of subdivided elements or they would merely select the individual subject heading elements and let a computer program build the subdivided strings. A standardized order of subdivision elements could also lead to a reduction in the number of errors that are due to subdivision order (Drabenstott and Vizine-Goetz 1994).

We do not believe that more studies would reveal different degrees of understanding among users. There is sufficient basis on which to determine that the order of subdivisions could be standardized without great loss of meaning.

If members of the library community are disturbed at the low levels of end-user understanding described here, then they should begin to consider making more drastic changes, possibly, along the lines suggested by Cochrane (1984). She recommended breaking up long subdivided subject headings and defended her recommendation saying that "the logic behind the string's construction is lost on most catalog users" (Cochrane 1986, 62). Now we have empirical evidence to support her claim.

If the library community does not make changes to the existing system, we have other recommendations that LC could introduce to improve end-user understanding. LC should consider involving people who are heavy users of the system—children, adults, and reference librarians—in the establishment of new subject headings and subdivisions in the LCSH system. Several types of consultation are possible. LC could sponsor clubs, committees, working groups, etc., of children and adults who would serve in an advisory capacity to the Cataloging Policy and Support Office (the editorial board for LCSH). Although children and adults could be recruited from nearby high schools or public libraries or could be frequent public users that LC's reference librarians have come to know, they could be recruited from afar and use electronic mail or other collaboration technologies to assist staff of the Cataloging Policy and Support Office. Also such staff could review published material on a subject across several different intended audiences to find language shared by audiences to express subjects. Surely staff of the Cataloging Policy and Support Office would have their own ideas about recruitment. It is important to include library catalog users—children, adults, and reference librarians—in the process.

Future researchers could experiment with introducing certain indicators to subject headings that would reduce the
problems library patrons have understanding subject headings due to syntax. Unfortunately, librarians would have to teach patrons how such indicators worked. They would be unable to reach all patrons to explain the system, and those patrons they did reach would probably forget the explanation rather quickly.

The punctuation between subject heading elements could be examined. Most catalogs combine such elements using two hyphens (--) or an em dash (—). Researchers could experiment with colons (:), slashes (/), or tildes (~) between elements. In recent years, the principal researcher of this study has observed new students or faculty colleagues in related fields explain how the LC subject subdivisions system works. They almost invariably describe a system in which the individual elements in subdivided subject headings are ordered in a hierarchical relationship. Perhaps researchers would consider undertaking studies that introduce different punctuation between subject heading elements to determine what impact such elements have on subject heading understanding.

CONCLUSION

It is time for the library community to grapple with difficult questions about its subject-access system and make informed decisions about solving the problem of low levels of end-user subject heading understanding. Subject analysis and its representation by LCSH have been the primary means of subject access in library catalogs for more than one hundred years. Practical reasons—including the enormity of the investment in this system—suggest it will continue into the next century. The findings of research into end-user understanding is an important source of information that can assist LC and subject catalogers in making decisions that contribute to the effectiveness of cataloging and the subject headings.

WORKS CITED


Notes on Operations

Spelling Errors in the Database: Shadow or Substance?

Barbara Nichols Randall

The purpose of this research was to determine the extent of spelling errors in the University at Albany's online catalog, whether these errors seriously affect users' access to library materials, and what effect spelling errors will have on the group database planned for the State University of New York. Using standard database tests, I studied the catalogs of the four University Centers (Albany, Binghamton, Buffalo, and Stony Brook) as well as two comparison catalogs: the New York State Library's Excelsior and the University of California's Melvyl. The results of these studies show that misspellings are unavoidable due to the way that most catalogs were built. These errors, however, are rarely an impediment to retrieval. I conclude with suggested ways to find and correct misspellings without expensive large-scale efforts.

A typographical error in a ship mortgage prepared by Haight, Gardner, Poor & Havens could cost the Prudential Insurance Co. of America between $11 million and $31.5 million before a dispute in federal court is finally resolved. . . . at issue is a $92.8 million lien. . . . three zeroes were dropped from the amount when the mortgage was amended in April 1986, leaving Prudential with a lien that may be worth only $92,885.—Frost and Goldner (1988, 7).

Damn construction 92 percent complete at Brushy Creek—Leno (1998).

Misspellings and typographical errors in library databases are neither as costly as those made in financial documents nor as funny as those highlighted every Monday night by Jay Leno on The Tonight Show. A literature review shows that misspellings and typographical errors have been, however, the subject of much research. Bourne (1977), Ryans (1978), Dwyer (1991), Ballard and Lifshin (1992), Gardner (1992), and Cahn (1994) all deal with the identification or effect of misspelling on a database. Bourne (1977) concentrated on misspellings of index terms, including the number of misspelled terms in computer databases, the implications of these misspellings to searchers, and who should clean up the errors. He found the occurrence of misspelled terms ranging from 1 in 8,000 citations in one database to 1 in 160 citations in another.

Ryans (1978) studied the accuracy of 700 records in the OCLC Online Computer Library Corporation, Inc., database...
using Anglo-American Cataloguing Rules (AACR), International Standard Bibliographic Description (Monographs) (ISBD (M)), and OCLC input standards as measures. Most of the errors she found were “due to simple carelessness” (131). She found errors on 283 records, including misspellings and typing errors. She did not quantify the errors, but she did describe them as “frequent.”

At the time of these early studies, online libraries were a dream rather than a reality; libraries throughout the country joined shared cataloging networks and began the preliminary work for the eventual computerization of their catalogs. In New York state, a number of large research libraries with adequate money and institutional computer expertise created their own online catalogs. Throughout the 1970s and 1980s, librarians performed extensive retrospective conversion of their card catalogs (Reed-Scott 1985). Federal, state, and local funds were used throughout the country, and catalog records were created by contractors and in-house catalogers, and through national cooperative projects such as the COMARC (Cooperative Machine Readable Cataloging) and the CONSER (CONversion of SERials) projects. Catalogs of varying quality often resulted, which led to projects to clean up the data.

Beall (AL Aside 1991) started a dialogue on misspelling that is ongoing, with almost quarterly discussions occurring on AUTOCAT. Beall searched the occurrence of 10 common word misspellings, totaled the occurrence of the words, eliminated the i.e. or sic words, subtracted the total from 100, and compared the result to other libraries of similar size. The 10 misspellings are: February, Guatemala, Mission, Government, Fransisco, Grammar, Receive, Wednesday, Separate, Conditions. Dwyer (1991) refined the method by deriving a way to measure a meaningful error rate by comparing the number of misspellings to the number of correct spellings of the words. Cahn (1994) used a measure to take into account whether access was prevented because of the uniqueness of the error or whether the error was redundant and therefore did not affect access to the record. Ballard (1992) published a list of commonly misspelled words in online catalogs as a result of a project begun in 1991 to rid the Adelphi University database of obvious typographical errors.

What does all this discussion of misspelling and typographical errors mean? Does the discussion represent merely the perspective of good spellers and proofreaders? Are our catalogs so flawed that our patrons won’t find what they need? According to psychologist Craig Brod (1984, 15), “Unwittingly, we are adopting as our own the computer’s standards. We have come to expect from people the perfection, accuracy, and speed to which computers have made us accustomed.” As librarians, we must attempt to separate the substance—errors that deny access to information—from the shadow—machine-like perfection.

**Purpose**

This study was conducted to determine, first, how dirty the University at Albany’s catalog data are, and second, the effect that these data will have on a group database planned for the State University of New York libraries (SUNYConnect). “Dirty” was defined here to have two meanings: first, the number of misspellings that occur in the database; and second, the degree to which misspellings inhibit access to the library’s materials.

**Method**

First, the research team searched the University at Albany’s catalog using the list of words from Beall (AL Aside 1991). We performed keyword searches and computed a Beall score to get the frequency of error. We next calculated the error rate defined by Dwyer (1991). Then we compared both values to the values found at the two other university libraries in our proposed group database that have similar-sized collections (based on self-reported data from the 1998–99 American Library Directory)—Binghamton and Stony Brook. After this preliminary comparison, we searched a set of words in one subject area (economics) both to locate misspellings and to determine
the relative importance of the errors found. Finally, the error rates at the fourth library in the proposed group were compared to the rates at the New York State Library, which has a collection of a similar size.

**RESULTS**

Table 1 presents the Beall scores for the three institutions. Table 2 presents the Dwyer scores for these same institutions. None of the numbers seem to be consequential, but because an important part of database maintenance is correcting errors, the question can be raised about when such errors might be ignored. The answer lies both in the placement of the error—that is, whether the error denies access or whether it does not—and in the uniqueness of the term in the record. To address this, the second stage of our research involved searching for variations on three words related to one subject: economy, economic, and economics. All misspelled variations of the words were found in the Albany, Binghamton, and Stony Brook databases. We found 16 variations in misspelling (see table 3): Albany had 12 spelling variations,

<table>
<thead>
<tr>
<th></th>
<th>Albany</th>
<th>Binghamton</th>
<th>Stony Brook</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors</td>
<td>Total</td>
<td>Rate</td>
<td>Errors</td>
</tr>
<tr>
<td>February</td>
<td>2</td>
<td>10,545</td>
<td>5,272.5</td>
</tr>
<tr>
<td>Guatemala</td>
<td>5</td>
<td>1,005</td>
<td>201</td>
</tr>
<tr>
<td>Mission</td>
<td>0</td>
<td>1,298</td>
<td>0</td>
</tr>
<tr>
<td>Government</td>
<td>9</td>
<td>62,640</td>
<td>6,960</td>
</tr>
<tr>
<td>Fransisco</td>
<td>4</td>
<td>5,078</td>
<td>1,260.5</td>
</tr>
<tr>
<td>Grammer</td>
<td>11</td>
<td>5,189</td>
<td>471.7</td>
</tr>
<tr>
<td>Recieve</td>
<td>1</td>
<td>145</td>
<td>145</td>
</tr>
<tr>
<td>Wensday</td>
<td>0</td>
<td>416</td>
<td>0</td>
</tr>
<tr>
<td>Seperate</td>
<td>22</td>
<td>1,964</td>
<td>89.3</td>
</tr>
<tr>
<td>Conditons</td>
<td>6</td>
<td>28,488</td>
<td>4,748</td>
</tr>
</tbody>
</table>

**TABLE 1**

**FREQUENCY OF ERRORS FOR TERMS IN BEALL'S LIST**

<table>
<thead>
<tr>
<th></th>
<th>Albany</th>
<th>Binghamton</th>
<th>Stony Brook</th>
</tr>
</thead>
<tbody>
<tr>
<td>February</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Guatemala</td>
<td>5</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Mission</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Government</td>
<td>9</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Fransisco</td>
<td>4</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Grammer</td>
<td>11</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Recieve</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Wensday</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Seperate</td>
<td>22</td>
<td>23</td>
<td>11</td>
</tr>
<tr>
<td>Conditons</td>
<td>6</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

**TABLE 2**

**DWAYER’S RATIOS**
Binghamton had 13, and Stony Brook had 9. The same variations in the same bibliographic records occurred in more than one database in 11 instances.

The 16 misspellings all were typographical errors—that is, errors in transcription, not misspellings in the original version. Typographical errors are not solely a byproduct of the computer age, but in fact have existed since early manuscripts were copied letter by letter in monasteries. Alfred Watts, a "printer's reader" (proofreader), wrote a classic work on typographical errors in 1883. Smith (1985) found Watts's work useful in understanding how to improve data entry and proofreading. Smith noted that Watts evaluated each type of typographical error found in a sample of 60 two-column pages of small type set by six different compositors. Watts classified errors into three categories: errors of omission, substitutions, and doubling. Gardner (1992) further refined the substitution and omission categories by extracting two additional categories: errors of letter transposition and errors of letter insertion. An error of transposition occurs when two adjacent letters are interchanged. Errors of insertion occur when an extra letter, either the same or different, is added to the word. Ballard and Lifshin (1992) identified typographical errors as errors of omission, substitution, insertion, transposition, added space, and dropped space.

Using Watts's error categories, the sample contained 6 instances of omission, 8 of substitution, and 2 of doubling in the sample (see table 4). When we consider Gardner's modifications, two of the substitution errors could be called errors of letter transposition and one of the two doubling errors could be called an insertion error. Finally, we did not conduct tests for Ballard and Lifshin's dropped space or added space errors. Spacing problems can be found in some catalogs using forms of internal truncation. However, we did not pursue this approach due to the uncertain results that would be achieved in Albany's database. It is possible that there are un-

<table>
<thead>
<tr>
<th>Term</th>
<th>Albany</th>
<th>Binghamton</th>
<th>Stony Brook</th>
</tr>
</thead>
<tbody>
<tr>
<td>economy</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>economic</td>
<td>6</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>econonic</td>
<td>7</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>economic</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>economc</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>econominc</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>economic</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>economnic</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>economic</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>economics</td>
<td>0</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>economics</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>economcs</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>economics</td>
<td>2</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>economcs</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>economcs</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>economcs</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

detected instances of both within the databases. As Smith (1985, 189) said, "Although new technology presents new pitfalls for compositors and proofreaders, the old ones—the ones caused by human imperfection—remain to humble us."

We identified unique errors as such if the misspelled word occurred only once in the record. Following Cahn's (1994) definition

<table>
<thead>
<tr>
<th>Omissions</th>
<th>Substitutions</th>
<th>Doublings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic</td>
<td>Economy</td>
<td>Economic*</td>
</tr>
<tr>
<td>Econmic</td>
<td>Economic</td>
<td>Economnic</td>
</tr>
<tr>
<td>Economcs</td>
<td>Econonic</td>
<td></td>
</tr>
<tr>
<td>Economics</td>
<td>Economics</td>
<td></td>
</tr>
<tr>
<td>Ecomomics</td>
<td>Ecomomics</td>
<td></td>
</tr>
<tr>
<td>*also insertion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>** also transposition</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
of redundancy, if the misspelled word appeared correctly spelled in another place in the record we classified it as redundant. Albany’s database had 40% unique errors (12 of 30) and the remaining errors were redundant. Binghamton’s database had 18% unique errors (9 of 49); while Stony Brook had 43% unique errors (10 of 23) (see table 5).

Unique errors are not all equal. For example, errors in a title or subject field are more serious than errors in a note. In a study of online catalog use for the Council on Library Resources, Larson (1983) concluded that most users search by subject. Anderson (1995) reiterates that users rely on keyword and subject searching to find information. Ballard and Lifshin (1992) found the majority of the errors in their study in title fields (63%), followed by note fields (21%), author errors (9%), and series errors (7%). The errors we found in this study occurred in five field types: author fields (including main author, alternate author, and publisher), title fields (including main and alternate titles), subject fields, note fields, and series fields. By analyzing the unique errors, we found that the majority of errors occurred in note fields, with title fields taking second place. There were no errors in subject fields. Redundant errors also followed this pattern.

As we enter the new century, the importance of regional or virtual catalogs has grown rapidly, which has implications for database errors. After the cleanup of an individual catalog, will the creation of a group catalog bring back the errors or compound the errors of our individual catalogs? We wanted to know the overlap of common misspellings in the Albany, Binghamton, and Stony Brook databases. Six of the misspellings occurred in all three databases; 5 of the 6 misspelled words had common records for more than one institution (see table 6). One record was common to all three databases. Eight records were common in two databases. The overlap of some of the typos was bothersome. We wanted to know whether these errors were all from data entry or whether some were the result of a common record that had errors. To check for this in the databases of the three institutions, we used the subset of economics records that existed in more than one database.

In addition to the common misspelled records, common records where one library corrected the database misspelling also existed. Of the total of 16 misspelled economics words, a total of 85 records had misspellings, and 23 records did not have the misspelling. Of the 17 common misspellings (two or more) in table 5, we found 4 records with the terms spelled correctly.

The next step was to look at the catalog of the fourth university center in the proposed group, Buffalo. We ran the Beall, Dwyer, and misspelled economics terms tests on Buffalo’s catalog. The Beall score was -40, and the Dwyer scores were considerably lower. The economics terms test revealed 12 misspellings. Of these, 1 was a misspelling not previously identified; 4 were misspellings also found in one other catalog; 2 were misspellings found in two other catalogs; and the remaining 5 were misspellings found in three other catalogs. We identified 3 additional records as common records and found 1 additional misspelling (economy).

### Table 5

<table>
<thead>
<tr>
<th>Database</th>
<th>Error Type</th>
<th>Author</th>
<th>Title</th>
<th>Subject</th>
<th>Note</th>
<th>Series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albany</td>
<td>Unique</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Redundant</td>
<td>3</td>
<td>11</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Binghamton</td>
<td>Unique</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Redundant</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>Stony Brook</td>
<td>Unique</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Redundant</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>
But we were unclear how to evaluate the scores found for Buffalo's database. We were unsure the effect that the size of the collection might have on the results, and thus the relationship between the Beall score of -40 found for Buffalo, and the scores for the other three institutions in the group. As a comparison, we compared Buffalo's scores to the New York State Library, which has a collection of similar size.

The State Library scored even lower on the Beall test, -60, and consistently lower on the Dwyer test. The economics terms test revealed 25 misspellings (see table 7). In comparison to the New York State Library, the quality of Buffalo's database was good.

Why is there such a difference between the State Library and Buffalo? Because it was possible to determine the source of the State Library’s records, but not those of the other institutions, those records were studied in depth. The State Library’s automated catalog, Excelsior, is a second-generation database. The original database, CMS (Collection Management System), was first operational in 1978 when the State Library moved its primary collections and base of operations from the Education Building to the Cultural Education Center in the Nelson E. Rockefeller Empire State Plaza, the seat of New York State government. The initial retrospective conversion work was done through a contract with a local nonlibrary contractor, Finserv. The State Library was also an early member of OCLC, so OCLC archive tapes were also used. The State Library is one of the original CONSER participants. Additional retrocon projects were performed throughout the 1980s, including cataloging the American Periodical Series, a grant-funded project to catalog the Goldsmiths’-Kress Library of Economic Literature, two contracted upgrade projects (one again through Finserv and the other through OCLC machine match), and the purchase and direct loading of the SuDocs (Superintendent of Documents) tapes as a full government documents depository. In descending order, the errors originated with Finserv (20), OCLC (19), SuDocs (13), OCLC upgrade (11), Goldsmiths’-Kress (8), Finserv upgrade (6), CONSER (3), in-house direct input (3), access level document cataloging (DACS) (2), and archival records from the State Archives and Records Administration (SARA) (2).

The majority of the errors occurred during retrospective conversion. This is not surprising given that one of the goals of retrocon is always production. All of the titles in the Finserv, Finserv upgrade, and

---

<table>
<thead>
<tr>
<th>Error</th>
<th>Title</th>
<th>Date</th>
<th>Databases</th>
</tr>
</thead>
<tbody>
<tr>
<td>economic</td>
<td>Economic progress</td>
<td>1955</td>
<td>Albany, Binghamton</td>
</tr>
<tr>
<td>economic</td>
<td>A high-speed passenger rail system for the U.S.</td>
<td>1981</td>
<td>Albany, Binghamton</td>
</tr>
<tr>
<td>economic</td>
<td>The economics of direct employment</td>
<td>1900</td>
<td>Binghamton, Stony Brook</td>
</tr>
<tr>
<td>economic</td>
<td>Interest as a source of personal income and tax revenue</td>
<td>1955</td>
<td>Albany, Stony Brook</td>
</tr>
<tr>
<td>economic</td>
<td>Our emergent civilization</td>
<td>1947</td>
<td>Albany, Binghamton, Stony Brook</td>
</tr>
<tr>
<td>economic</td>
<td>Miscellaneous essays and addresses</td>
<td>1904</td>
<td>Albany, Binghamton</td>
</tr>
<tr>
<td>economics</td>
<td>Philosophy of economics</td>
<td>1982</td>
<td>Binghamton, Stony Brook</td>
</tr>
<tr>
<td>economics</td>
<td>Applied economic forecasting</td>
<td>1971/1966</td>
<td>Binghamton, Stony Brook</td>
</tr>
<tr>
<td>economics</td>
<td>Teachers as agents of national development</td>
<td>1971</td>
<td>Binghamton, Stony Brook</td>
</tr>
</tbody>
</table>
OCLC upgrade groups were older material matched or upgraded based on shelflist cards. The Finserv project was a separate, production-oriented project undertaken while the move to the new library building was occurring. The library's cataloging staff were not involved in the project initially. The Finserv upgrade and OCLC upgrade projects did have cataloging staff involvement as well as extensive systems evaluation. The inclusion of CONSER records, which underwent rigorous review, in the error group illustrates the very human nature of spelling errors.

OCLC is the bibliographic utility the State Library uses. We wondered whether the errors from the OCLC and OCLC upgrade records still existed in the OCLC database. Thirteen of the 30 records remain misspelled on the OCLC database. This percentage rate, 43%, is lower than the 51% documented by Ballard and Lifshin (1992). Three were in fields input by State Library staff at the time of production and were never present in the OCLC database.

We searched one last database: the University of California (UC) union catalog, Melvyl. Melvyl is a model for the SUNYConnect project. The Melvyl Union Catalog is part of a statewide computer-based library system created in 1981 by the California Digital Library (formerly known as the Division of Library Automation), in conjunction with UC campuses. It has been available online since the mid-1980s (Crowell 1995) and has been available in Web format since 1997. We wanted to compare the error rates found in this almost twenty-year-old catalog.

We searched the Web version of Melvyl using title keyword, subject keyword, personal and corporate author keyword, and series keyword. The note field is not keyword searchable in Melvyl. Melvyl (with 9,678,014 titles and 14,632,800 holdings as of November 25, 1998) is approximately twice the size of the combined university centers, discounting overlap. We found error frequencies for the Beall and the economics term tests. We could not run the Dwyer test because Melvyl does not allow for complete counting of the correctly spelled words. Any Melvyl search that retrieves more than 10,000 hits is stopped.

We found at least one instance of each misspelling. The primary category of errors was subject fields, followed by title, author, and series. Seven of the nine overlap titles in table 6 were also owned by UC libraries and included the same errors.

Finally, due to the way that data entry production is measured, we broke down typographical errors as error per character input, whether written or typed. Watts (Smith 1985) found 1 error in every 1,750 characters. Chan (1994) notes that almost one hundred years later, in July 1980, researchers at the National Composition Association found an error rate of 1 in 1,000.

### TABLE 7

**ERRORS IN ECONOMICS TERMS**

<table>
<thead>
<tr>
<th>Term</th>
<th>Buffalo</th>
<th>NY State Library</th>
</tr>
</thead>
<tbody>
<tr>
<td>ecomomy</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>economy</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>economy</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>economic</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>economics</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>econominc</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>economic</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>economic</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>economic</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>economc</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>econimic</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>economonic</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>econommmic</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>economnic</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>economnic</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>economc</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>economio</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>economics</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>economics</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>econonics</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>econimies</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>econmics</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>economcs</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>economcs</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>econoomics</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
The University at Albany’s error rate is 1 in 1,946 characters.

CONCLUSION

While perfection in both humans and databases is a worthwhile goal, the reality is that it is also impossible. Most spelling errors are redundant errors and thus, do not prevent users from finding the needed record. As Cahn (1994, 30) correctly stated, “Issues of time and money cannot be ignored.” Bourne (1977, 9-10) called misspellings “internal parasites to the search system,” yet he said, “while a relatively large number of index terms are misspelled (compared to conventional printed indexes), and while those errors are very conspicuous, they in fact have relatively little impact of file use for many of the databases.” Most librarians probably will agree with Ballard and Lifshin (1992, 139), who pointed out, “It may be widely perceived that spelling errors in OPACs and other large databases are few in number, randomly distributed, and impossible to locate in any systematic fashion. . . . every library that has an OPAC with keyword capability should search the problem words that we have identified and fix the inevitable errors.”

In consultation with the head of cataloging at the University Libraries, we chose to take a staged approach to database cleanup. Although the consensus was that the errors were minor, we wanted to search for the terms in Ballard’s list. Because we lacked staff time and money, a volunteer conducted the search for us. The volunteer noted only the number of occurrences of the misspelled terms and found 697 potential misspellings of 106 words on the Ballard list. Some misspellings were in fact correct transcriptions of title page errors and on investigation were correctly labeled “i.e.” or “[sic].” There were no instances of misspellings for 73 of the words found on the Ballard list. This list is being used as a guide for correction as staff or students become available to do the work. We are correcting the most frequently misspelled words first.

Concern about the impact of misspellings on the catalog should be minor. Although misspellings or typos are embarrassing, the low number of unique occurrences of each misspelled term means that in most cases a user will still be able to find a relevant item. The breakdown of the State Library’s misspelled records suggest that most of the misspellings or typos might be traced to the profession’s early years of retrospective conversion. Given that this work can be sorted out and listed by project code, it becomes possible to target these records for further examination. Moreover, some of those records will disappear as we inventory and weed our collections.

Reports from our systems people bode well for our quest for perfection. Each time we request a report to work on an area of problems in the catalog, we also find a few other problems, usually misspellings. Recently, when requesting a report of all the unbracketed general media designators in the title transcription area, we also generated a list of instances of “microform,” “midroform,” and “video recording.” We are aided in our perfection quest by sharp-eyed spellers who send our department errors they notice in the catalog.

However, correcting spelling errors and typos can take more time than most libraries have. The small number of these errors that can be reduced further (given the concept of uniqueness) shows that the type of large-scale effort Ballard (1992) performed at Adelphi University is beyond the means and needs of most libraries. The combination of errors that occur as union catalogs, whether virtual or otherwise, are created increases the number of errors but not by any consequential amount. We believe that our current error-correction efforts, on an as-needed basis or as a byproduct of other enhancement projects, are both sufficient and reasonable.

WORKS CITED


Ballard, Terry. 1992. Spelling and typographical errors in library databases: One library’s
system for rooting out spelling errors. Computers in libraries 12, no. 6: 14–19.
Combating Whole-Book Deterioration: The Rebinding & Mass Deacidification Program at the Penn State University Libraries

L. Suzanne Kellerman

In this article, I describe the newly established preservation program at the Penn State University Libraries to deacidify en masse collections of older materials en route for commercial rebinding. This collection-based preservation strategy was designed to serve as a one-step-preparation approach for rebinding and for neutralizing the acid paper found in our heavily used, worn, and torn circulating collection. This strategy of combining commercial rebinding with deacidification has resulted in a timely and effective solution to preserve and stabilize collections for our users now while preventing future material deterioration. I will detail agreements and operational procedures developed among service vendors and the library, describe selection criteria, and document technical aspects to report this preservation action in bibliographic records.

Since the 1960s, librarians and collection managers intently have watched the rapid evolution and development of mass deacidification technology as services, products, and vendors have come and gone over time (Boomgaarden 1995; Buchanan 1994; Cloonan 1990). The mass deacidification process, a preventive option to preserve acidic but-not-yet brittle and brittle materials in their original format, was once considered an unsafe and unreliable process. Today, this public perception and opinion has changed as researchers have shown in many scientific tests and studies that some processes indeed are effective in combating acidic paper degradation (Evaluation 1994). Additionally, when the Library of Congress conducted its own research and testing, and then awarded its first mass deacidification contract to Preservation Technologies, Inc. in 1996, many librarians began to reconsider mass deacidification seriously as a viable preservation tool (Harris 1998). At least one vendor reports that subsequent to these developments, contractual agreements between mass-deacidification vendors and libraries have been rising.

BACKGROUND

Like many who were interested in the prospects of mass deacidification, librarians at the Penn State University Libraries (PSUL) began in 1991 to investigate vari-
ous deacidification processes as a means to test claims that these techniques could safeguard paper-based materials printed on acidic paper (De Stefano 1994; Mass deacidification 1992; Sparks 1992). We conducted several pilot projects in 1991 and 1992 with the top three vendors of the time: FMC Lithium Corporation of America (Lithco), Akzo Chemical Inc., and Preservation Technologies, Inc. (PTI). Results of these independent tests were mixed. Our findings showed that some treatments adversely affected some types of books, papers, inks, and labels, while others displayed no negative effects. We were most encouraged by PTI’s Bookkeeper process.

From 1992 to 1995, librarians worked with PTI to test enhanced treatment processes that PTI had developed and to establish in-house operational procedures and selection criteria policies. It soon became evident that books needed to be in good structural condition and fully intact to endure the treatment process. We discovered from these early tests that books in poor physical condition, e.g., volumes with loose boards or paper covers, broken spines, or detached pages, could not easily withstand the rigors of the treatment procedure and were further damaged by the mechanical agitation inherent in the process. In addition, books selected for the Bookkeeper treatment had to be durable and fully flexible to be attached to metal flanges splayed out or fanned open, which allowed each page and inner gutter margin to be exposed fully. This permitted a complete and even distribution of the neutralizing agent to all sections of the book. From these early tests, it became immediately apparent that books in poor physical condition would first require structural stabilization (first time binding or rebinding) to survive the treatment process successfully.

**THE INITIAL PROJECT**

In 1995, PSUL formally contracted its first mass deacidification project with PTI. We selected 163 volumes from among 2,000 volumes in the African American Studies Room for this initial project. Due to limited funding, we applied certain selection criteria to identify the volumes that were the best candidates for the process. We assessed books individually to determine structural integrity and level of paper acidity. Only structurally sound books printed on acidic paper made the first cut as potential deacidification candidates. We further narrowed this subcollection based on circulation use patterns, giving a priority to the most heavily used items, or those with a circulation count of 8 or more. Books meeting all of the above criteria had to be physically available, i.e., on the shelf at the time of the assessment.

We preselected materials over a two-month period and sent them to PTI in June 1995. A month later, treated books were returned to the library, where we conducted physical inspection and random pH testing (using a pH testing pen) on them. Results of the testing showed that 85% of the deacidified materials (138 books) had been treated successfully with no adverse effects to papers, inks, or labels. Of the remaining 15% (25 books), the major problem was that of partial treatment (nonuniformity), especially in the inner margins of some volumes (Evaluation 1994).

The success of this first, small-scale mass deacidification project encouraged us to plan an ongoing, larger scale mass deacidification program for PSUL. We continued to work with PTI to resolve the “fish-eye” problem of partial treatment found in some materials and focused our attention on formulating a plan to implement this new preservation program.

Based on our experience working with PTI’s Bookkeeper process, we learned several important factors that influenced the success or failure of using this treatment process as a viable preservation option. First, books had to be structurally sound to withstand the rigors of the process. Second, not all paper types absorbed the neutralizing agent equally well, e.g., clay-coated paper. Third, tracking each item sent for treatment was essential in order to keep library patrons informed of material availability. Fourth, internally, we recognized the need to implement the use of the USMARC 583 field to report
this preservation action in our local bibliographic records for our users. Fifth, we needed to identify an ongoing funding source to support the program. Finally, we needed to develop selection criteria to determine which library collections or portions of collections were best suited for mass deacidification.

**SELECTION CRITERIA FOR MASS DEACIDIFICATION**

At PSUL, we initially planned to target portions of our older collections and some new acquisitions for mass deacidification treatment based on the findings of a study we conducted in 1992. During this 1992 study, we surveyed and tested 5,362 new books with imprint dates of 1990, 1991, and 1992, received through our approval plan programs, to determine the percentage of books printed on acidic paper versus those printed on alkaline paper. Results of the study revealed that of all books surveyed (hardbacks and paperbacks from 33 countries worldwide), 86% were printed on alkaline paper; only 14% were printed on acidic paper. Of the books published in the U.S., 91% were printed on acid-free paper. Although impressive, these figures were not all that surprising in light of the 1989–90 call for commercial publishers to use permanent paper and for paper mills to produce acid-free paper. Production of alkaline paper was becoming the norm in paper production in North America (*Preserving Knowledge* 1990). Books published in third-world regions were the exception, e.g., books published in India and Africa were highly acidic. With only one exception, all books from these two countries were printed on acidic paper.

Based on our study, it became apparent that a high percentage of newly acquired materials, printed on alkaline paper, were not at risk as previously thought. Our older collections residing in the stacks were more likely to have been printed on acidic paper, were more subject to deterioration, and therefore were at far greater risk. As noted above, new materials published in some third-world countries were exceptions to this rule.

In 1997, five years after our initial collection survey, we conducted two additional studies to determine whether any changes had occurred in the publishing industry that might affect our earlier decision about which collections would benefit most from deacidification treatment. We surveyed newly acquired materials and older circulated materials sent to the Preservation Collections Care Unit for conservation treatment (repair, rebinding).

We performed a pH test on 447 newly acquired books (both paperbacks and publisher bound materials) to determine paper acidity. Domestic and non-U.S. published materials received through firm orders, approval plans, or monograph and serial acquisitions processes were represented among these titles. Of the 447 volumes tested, less than 1% (.06%) were found to be printed on acidic paper.

Later that year, over a two-month period in the fall semester, we conducted a pH test on 156 recently circulated books returned from patron use. As books were returned to Lending Services, staff identified materials in poor physical condition and sent them to the Preservation Collections Care Unit for repair. Condition of these materials varied and included loose or missing covers, torn or loose pages, broken spines or headcaps, and missing pages. These 156 volumes represented publications from both the U.S. (96%) and abroad (4%). Of the 156 volumes surveyed, we found that 82% were printed on acidic paper; 18% were printed on alkaline paper.

Based on the survey results from these two tests, it was apparent that the older materials were most at risk by evidence of their poor worn and torn physical condition and their acid content. We then created a one-step-preparation approach for combining rebinding with mass deacidification as a whole-book strategy. Previously, other libraries using mass deacidification had used similar selection criteria for identifying suitable candidates, but varied in the sequence of the treatment process (De Stefano 1994).

In April 1998, we held a meeting with library selectors to share these survey test results and to discuss the combined re-
binding and mass deacidification program as a treatment option for acidic collections. Selectors at the meeting agreed that the combined process was a viable preventive preservation option and that the following materials should be considered for treatment, listed in order of priority:

1. Circulating materials needing conservation attention, as identified by Lending Service staff (from the University Park main campus location only)
2. Selected discreet collections of distinction identified by library selectors
3. Selected individual titles identified by library selectors

Once we established these selection priorities, librarians in the Preservation Department accelerated their efforts to finalize planning of operational procedures for a combined rebinding and mass deacidification program for PSUL. In addition, a newly established preservation endowment secured financing to support this new preservation program. We held numerous collaborative meetings with the service vendors involved (both a commercial binder and a mass deacidification service provider) to establish service agreements. We also held meetings with staff from various library departments (Lending Services, Stack Maintenance, Cataloging and the Business Office) to establish internal workflow procedures for handling, tracking, shipping, and billing.

PSUL established the agreements and operational procedures presented in this paper in collaboration with its vendors. We have outlined them here as an example of how similar projects might be structured. The first section describes mutually agreed-upon library requirements and vendor specifications for the process. The following section documents the internal operating procedures established for sending worn and torn circulating materials for rebinding and mass deacidification treatments.

LIBRARY REQUIREMENTS AND VENDOR SPECIFICATIONS

Having committed to this unique one-step preparation process as an acceptable preservation option, PSUL established basic preservation requirements for the rebinding and mass deacidification treatment processes. One of these requirements was that all materials be rebound or first time bound using C-Grade bookcloth following the standards for commercial binding set forth in Parisi and Merrill-Oldham (1986). We also determined that the acceptable range of pH levels for treated materials would be 6.5 to 10.4 with a minimum required alkaline reserve level of 1%. Both vendors had to notify the library if materials could not be bound or mass deacidified for any reason. The binding vendor had to supply a shipping schedule. The deacidification vendor had to supply surrogate pH testing paper reports for each shipment and also track all incoming items by call number or barcode for easy identification.

In addition to our library requirements, we profiled vendor-specified stipulations to ensure efficient throughput and turnaround time. Working collaboratively, the library staff, binder, and deacidification vendor defined additional program specifications. The library staff agreed that each shipment would have a minimum number of 30 volumes and that items would not exceed 12" x 9" x 2.5" in size nor weigh less than 2.8 pounds. The binder required the library to send materials two weeks prior to their shipment to the mass deacidification vendor to ensure sufficient time for rebinding. The binder agreed to rebind and ship bound books to and from the deacidification vendor’s facility, and then transport treated books back to the library within a total turnaround time of 6 to 8 weeks.

It was also agreed that the binder would arrange the shipment from smallest to largest to facilitate processing and handling by the mass deacidification vendor. In turn, the deacidification vendor would return the books in the order they were received. Other program requirements called for the binder to contact the library if any materials were found not suitable for mass deacidification (e.g., volumes that were brittle, oversized, or that weighed more than 2.8 pounds), and for the deacidification vendor to supply
vendor identification labels stamped with the treatment date for in-house application to each treated book. Finally, each vendor had to submit separate invoices for products and services rendered.

Once an agreement with external vendors was reached, PSUL then initiated internal procedures to implement the process.

**INTERNAL LIBRARY OPERATIONS**

Based on the Preservation Collections Care Unit staff's current responsibilities of assessing materials for preservation treatment, it seemed logical for staff in this unit to assume responsibility for managing the mass deacidification program. While many of the assessment activities were similar to ones already in place, we needed to add some new procedures. These included evaluating and selecting materials based on newly established criteria, bibliographic processing, physical preparation, and online tracking of materials sent for rebinding or first-time binding and mass deacidification.

Following the established selection criteria, the staff conducted item-by-item preselection assessment to identify candidates for mass deacidification treatment. Physical and use-related factors used to identify candidates for mass deacidification treatment included:

- Any volume (bound or not bound) printed on acidic wood-pulp paper
- Paperbacks or hardbacks in poor condition and needing to be rebound prior to mass deacidification treatment
- Items limited in size up to and including 12" x 9" x 2.5" (vendor requirement)
- Volumes weighing 2.8 pounds or less (vendor requirement), and
- Items with a circulation count of 5 or more

Based on previous experience, internal decisions, and discussions with vendors, we identified the following materials and format types as not being appropriate for mass deacidification treatment:

- Text paper that is alkaline
- Manuscript materials
- Items published on glossy clay-coated paper
- Leather covers with red rot

- Materials with plasticized covers
- Brittle text paper (if fails at two double folds)
- Materials with accompanying pocket parts
- Titles that are not permanently retained
- Titles found on microfilm or microfiche

To determine whether published texts were printed on acidic paper, we performed a pH test. Using a pH chlorophenol red pen, we drew a small line on the tail of each volume closest to the spine. The chlorophenol red indicator solution turned purple if the pH of the paper was 6.8 or above, indicating that the paper was alkaline. A clear or yellow color marking indicated the paper was acidic and the volume was a candidate for mass deacidification. We chose this method for testing pH solely for its ease of use and process efficiency in determining a general acidity reading.

**BIBLIOGRAPHIC PROCESSING**

Once we assessed the materials, tested them for paper acidity, and sorted them for mass deacidification treatment, staff in Collections Care updated the bibliographic cataloging records for each title or volume selected. They updated item records and charged out each item to a single department charge, coded MASS. The automatic mnemonic display code MASS was established to appear in the online catalog to alert users and library staff that the item had been sent for deacidification, and was temporarily unavailable (see figure 1). This mnemonic code displays whenever the status of an item is requested through an “item inquiry” command. Or, if a patron checks the status of an item, the display will read “1 Dept. Charge,” also indicating the item is unavailable for use (see figure 2).

While the USMARC 583 standard has yet to be officially approved as the bibliographic field to record information about processing and preservation action taken on bibliographic materials, we adopted ALA recommended guidelines and terminology for using notes field 583 for this

HE5623.J47
-Central Pattee Level 4
(Location in stacks)Blue-1 Dept. Charge

Figure 2. Online Catalog Status Display Showing an Item that Has Been Department Charged.

was treated, a 598 note field was added to identify the specific copy that has been treated (see figure 3). The 598 field is defined as a local free text note used for reporting internal processing. The local definition of field 598 is:

\[ 598 \$a \text{Processing instructions} \]

Other examples of reporting when a specific volume was deacidified include:
- 598 Engineering copy deacidified
- 598 Physical Sciences v. 2 deacidified
- 598 [barcode #] deacidified
- 598 Stacks v.1 deacidified

**MATERIALS PREPARATION**

After we completed the online bibliographic updating, we generated a two-part printout of the bibliographic record (see figure 4). We detailed binding and spine stamping instructions as well as notification of mass deacidification treatment on the printout. We inserted a portion of the printout sheet into the volume being sent for treatment; we placed the extra sheet in a packet. We filed each packet alphabetically by title and then by shipping date.

We took and recorded separate statistics for volumes prepared for rebinding and mass deacidification. We then assembled and packed the books into a sturdy commercial binder shipping box marked “MASS DEACIDIFICATION,” and sent the box to the commercial binding vendor.

**POST-TREATMENT PROCESSING**

Once items were returned to the library, the Collections Care staff reviewed and inspected each volume for accuracy in rebinding (letter stamping, etc.) and for paper acidity. Staff used paperwork to compare requested binding type and spine stamping with the book in hand. Staff members also visually inspected each book for the proper placement of the infinity symbol at the top of the spine to identify that the book had received mass deacidification treatment. Adding this mark to the outside of each treated book allowed for a quick visual display of which books have and have not been treated. A mass deacidification vendor identification label on the verso of the title page
provided treatment information for library staff and patrons. The staff then conducted a random pH test to confirm treatment results.

After completing the physical and technical inspections, staff then added a giftplate to the book's inside front cover to denote the funding source used for the book's preservation. Staff then updated each item record and discharged the item, making the item available for use in the online catalog. Books were then ready to be reshelved and were returned to circulation.

As of this writing, we have prepared and shipped three 100-volume shipments for rebinding and mass deacidification and has received two return shipments. Turnaround time for each shipment was seven weeks. From all indications, the vendor and library specifications and requirements outlined have accurately described the process. Neither the vendors nor the library staff encountered problems in preparing or inspecting returned materials, so no changes in specifications

---

**Figure 3. Online Record Showing Use of Multiple 583 Fields and the 598 Field to Report Preservation Action Taken.**

**CONCLUSION**

As of this writing, we have prepared and shipped three 100-volume shipments for rebinding and mass deacidification and has received two return shipments. Turnaround time for each shipment was seven weeks. From all indications, the vendor and library specifications and requirements outlined have accurately described the process. Neither the vendors nor the library staff encountered problems in preparing or inspecting returned materials, so no changes in specifications.
or procedure were required. Although some fine-tuning of the procedures might be necessary as this program matures, for now it appears this new preservation strategy could become our most efficient process for stabilizing older collections using a one-step preparation approach.

WORKS CITED


INDEX TO ADVERTISERS

Library Technologies 2d cover
Archival Products 169
Haworth Press 3rd cover
OCLC 4th cover
Book Reviews

Margaret Rohdy, Editor


In this collection of essays on the impact of the current explosive development in information technology on academe, the editors make the case that academic librarians, educators, university administrators, and information technologists contributed papers to this volume. Because each essay stands on its own, smaller buildings, each useful in its own right, but with uncertain relationships with its neighbors. These factors might make reading this book difficult, but the work is nonetheless important, and one that warrants perseverance.

Those who write about the future of libraries commonly describe possible developments as inevitabilities and treat challenges as exciting opportunities. In contrast, the vision presented in The Mirage of Continuity is an inexorable and difficult future where the decision will not be whether we wish to alter practices, but instead, since we must change, how we can continue to fulfill our mission. The question is not how libraries must change, but whether libraries as discrete entities with their own collections, attached to individual universities, will continue to exist. The fact that the essayists do not describe these potential changes as fun commands attention. It is tempting to disregard predictions of wide-eyed techno-enthusiasts, but harder to dismiss this group of scholars who are remarkably unenthusiastic and give the impression of having come to their conclusions reluctantly. Their collective message, the breadth of concerns covered, and the uncanny agreement among papers are compelling.

In the first section, "Defining the Problem," the authors describe the nature of the changes facing academe as a result of the computer revolution. They assert that the incrementalist approach to change that is characteristic of both universities and libraries and the practice of viewing each development as another point along a continuum will no longer serve. Information distribution, retrieval, and use will be so different from the way they were when our current academic structures were developed that the future will be discontinuous from the past, and radical modifications of the academic enterprise will be required. To establish the context for these changes, this section includes a summary of the history of universities in the United States, the role and culture of the faculty, trends in higher education, and societal expectations.

The section titled "Integrating Information Resources with the Institutional Mission" is a provocative articulation of the purposes and responsibilities of
universities: producing knowledge, affording access to academic communities, providing and managing information resources, and serving as a means of representing accomplishment. It also is a pointed discussion of the patterns of scholarly communication and the interrelationships of faculty research, academic evaluation, publishing patterns, research libraries, and intellectual property rights. Emerging information technologies may both demand and enable restructuring of academe to assure that these purposes and responsibilities are satisfied. The structure and authoritativeness of library collections and the assumption by libraries of responsibility for preservation are contrasted to the relatively unstructured, uncontrolled, and non-preservationist nature of the Internet.

In the section “Challenges in Implementation,” after noting the increasingly prohibitive costs of materials, personnel, and space, Brian Hawkins states that “As great as the economic threats to libraries are, the greater threat is the perception that technology will solve these problems...” (129). He argues in favor of a new paradigm for collecting and providing access to information, involving de-institutionalization and collective remote data repositories. Richard N. Katz outlines assumptions underlying “the premise that academic information resources must be reconfigured in the first significant way since the opening of the Alexandrian Library” (155), and describes their implications, ranging from an imperative for collaboration among all segments of the university, to the need for standards and for rethinking intellectual property rights.

In the other papers, the essayists explore issues relating to creation, preservation of, and access to digital information resources, and posit that regardless of how or where information resources are held, current means of bibliographic control and information retrieval will be insufficient for the needs of users, though the Web may be even less satisfactory.

In “Leadership, Staffing and Management,” the future information resources professional is described as an “eclectic member of the university community, and a person who can span the boundaries of the various subunits on campus” (265), but in order for such persons to exist and to provide the kind of leadership necessary for steering a course through a transformational period, there will need to be substantive modifications in the preparation and mindset of these professionals, as well as changes in how such people are viewed in the university, and even in how successes are measured.

The Mirage of Continuity is an uncomfortable book to read. It is a success like few others in provoking thought and persuading readers that the digital revolution will bring changes that cannot be ignored until we retire and that cannot be handled by grafting solutions onto existing structures. The book’s particular strength is its breadth of vision. The essayists see the future in the context of academe as a whole rather than in terms of organizational details. Although libraries are central to their discussions, the authors do not treat them as self-contained institutions. Instead, libraries are considered as collections of functions, principles and purposes that have been located in a particular organizational unit, but which may be disaggregated as part of our response to the approaching discontinuity in the world of information resources.

This work is not a prescription for the future, and it is not a survival guide. It is instead a detailed, thoughtful, and compelling fair warning, or even a call to arms that academic librarians would do well to take to heart.—Janet Swan Hill (hillsj@spot.colorado.edu), University of Colorado Libraries, Boulder


Many of the authors presenting papers at the International Conference on the Principles and Future Development of AACR (Anglo-American Cataloguing
Rules) advocate moving cataloging boldly into the new millennium. Their recommendations address many of AACR's most significant weaknesses: the lack of a statement of principles, the absence of an explication of the rationale behind many of the rules, inconsistencies in treatment of content and carrier, and inadequate treatment of various types of materials, demonstrated most clearly by the lack of a general rule for new works of mixed responsibility and problems with the dichotomous monographs/serials model of the bibliographic universe. Because all of the papers are worthy of discussion, each will be discussed separately here.

In his introductory paper, "Modeling the Logic of AACR," Tom Delsey advocates the use of modeling techniques (e.g., entity-relationship modeling) to illuminate the structure and assumptions underlying AACR. As an example of how such modeling could be of assistance, Delsey tackles the dichotomy between intellectual content and physical form in the cataloging code, in which rules for description are supposedly based on physical form, and rules for access, on intellectual content. In an insightful and probing analysis of the rules reflecting this issue, Delsey discovers many contradictions and complexities. For instance, while some of the chapters in part I of the Anglo-American Cataloguing Rules, Second Edition (AACR2), base description on the physical form of items (e.g., sound recordings and videorecordings), others do not (e.g., cartographic materials and music). This paper provides a critical assessment essential for understanding AACR2's approach to form and content in the cataloging of an item.

In "AACR2 and Catalogue Production Technology," Rahmatollah Fattahi examines existing cataloging principles and concepts in light of their relevance in the online environment. Not surprisingly, he concludes that the finding and collocating functions of the catalog are still relevant, and suggests that cataloging practice be expanded to enhance functionality in the online environment, for example, to require additional descriptive elements such as tables of contents and summaries. Fattahi urges code makers to clarify the rationale for and functions of the concepts of main entry, uniform headings for titles and persons, and content and form of name headings in online cataloging environments. In addition, he believes that AACR2 should provide detailed guidelines for various levels of catalog displays as well as guidelines for the indexing of fields and subfields in online catalogs.

Martha M. Yee tackles many of the problems in AACR2's treatment of works in her paper, "What is a Work?" Yee makes more excellent points and suggestions in this paper than can be covered in a brief review, so I will discuss only a few of them here. Yee analyzes the criteria currently used in AACR2 to determine whether or not a change in an item justifies the creation of a new bibliographic record, finding them to be case-based as opposed to principle-based. She suggests a more principled approach using "fundamental content" such as text, music, and spatial data to determine the status of a particular item with respect to an existing work. She focuses on one of AACR2's most glaring weaknesses—the lack of general rules for entry of new works of mixed responsibility—and suggests how such rules might be created. Yee ends her paper with several pages of suggestions for changes in AACR2, in particular, changes to chapter 21, "Choice of Access Points." Like Fattahi, she urges including statements of objectives and principles to guide catalogers in decision-making.

Sherry L. Vellucci's "Bibliographic Relationships" is an excellent critical overview of the definition and treatment of bibliographic relationships in cataloging history. She ties together theoretical concerns, current practice in AACR2, MARC (MAchine Readable Cataloging), and online catalogs, and recent work on entity-relationship modeling, for example, the International Federation of Library Association's (IFLA) Functional Requirements for Bibliographic Records (1998), placing all in the context of user needs and research. Vellucci includes many suggestions for making the handling of bibliographic relationships in cataloging practice more consistent, concluding with
four general principles of bibliographic relationships to guide future revisions of the code: (1) relationship identification: bibliographic records should identify all important bibliographic relationships; (2) enabling linkage: data elements of bibliographic records should enable related bibliographic records to be linked; (3) multilevel description: the code should provide for description at several levels, including work, expression, physical item, and specific copy; and (4) consistency: identification and linkage of like relationships should be treated in a consistent manner.

In “Content versus Carrier,” Lynne C. Howarth provides a much-needed critique of the contradictions of a code that bases description on the physical form of an item (rule 0.24) and access points on its intellectual or artistic content (rule 20.1). Howarth notes the profession’s shift in emphasis from the creation of a surrogate based on carrier to one based on content. She advocates including both content and carrier into our view in a mutually inclusive way by implementing a four-tier model record based on recommendations from IFLA’s Functional Requirements for Bibliographic Records (1998) and from the Multiple Versions Forum Report (1990). In this model, a bibliographic record would include elements of description common to any work (“work level” tier); access points linked to authority records (“authority level” tier); unique physical properties or format-specific details (“manifestation level” tier); and copy specific information (“item-level” tier). While this model is not entirely without problems, it may have potential to improve access to resources described within the existing framework of AACR.

Michael Gorman and Pat Oddy review the history, principles, and impact of AACR2 in their “The Anglo-American Cataloguing Rules, Second Edition.” Perhaps their contribution—urging caution in integrating radical changes into AACR2—is de rigueur at a conference devoted to the future development of AACR, considering that many of the changes advocated in the other papers may be considered somewhat revolutionary. The authors make several modestly useful recommendations, including purging rules for special cases and resolving the problem of unpublished items.

“Issues Related to Seriality” by Jean Hirons and Crystal Graham is, in my opinion, the most significant contribution to cataloging theory of this decade. Hirons and Graham propose a new model of the bibliographic universe, consisting of static and ongoing publications (as opposed to monographs and serials), and make a series of intelligent recommendations for its incorporation into AACR2. Some of the more intriguing recommendations include adopting a three-dimensional approach to the cataloging rules, which would incorporate content, carrier, and publication status; creating rules for ongoing publications that focus on identification as opposed to transcription; and creating a new conceptual first chapter in AACR2 that would emphasize what the cataloger seeks to accomplish and why. This paper should be required reading in library schools (it will be in my classes!) and should also be read widely by practitioners.

In “Access Points for Works,” Ronald Hagler reviews the history of work identification in the cataloging code and in catalog technology, pointing out the importance of the main entry as a mechanism for identifying works. He also decries the optionality of uniform titles (chapter 25) in AACR2, calling it a “cop-out” (219). Hagler’s emphasis in this paper is on filing, browsing, and display considerations in online catalogs—the issues that affect users most. His recommendations to the Joint Steering Committee focus on work identification, and include changing terminology to distinguish “work” and “document” more clearly; placing AACR2’s context squarely in the online environment as opposed to the manual environment; and requiring catalog agencies to “provide access to every work appearing within each catalogued document” (227).

In “Beyond MARC” Mick Ridley pinpoints several of AACR2’s shortcomings, including the optional status of uniform titles and problems with the treatment of physical format. He is also critical of MARC’s proliferation in different versions,
as well as its difficulty in structuring information hierarchically. Ridley suggests that we move toward a “work-based system” that accommodates three hierarchical levels: work, manifestation, and copy. He also suggests that we move toward a more commonly used format such as an SGML application. Finally, he addresses the problem of character sets, urging adoption of Unicode to correct the inadequate display of language scripts in current systems.

This volume contains ideas and recommendations with which the cataloging community must acquaint itself to ensure that AACR2 does not stagnate, but becomes the rational, responsive, and flexible tool that it must be to sustain intelligent cataloging practice in an international context. The conference organizers are to be congratulated for selecting contributors who have produced works of such importance to the cataloging community.—Allyson Carlyle (acarlyle@washington.edu), School of Library and Information Science, University of Washington, Seattle.

Works Cited


Management of Serials in Libraries.
Serials always have been known for their complexity and for the many challenges they present to the people who manage them. Today, the challenges are no longer simply changes in title, frequency, and numbering schemes; there are also changes in format, methods of acquisition, and methods of delivery. Add the Internet, licensing contracts, and dwindling budgets, and the serials specialist faces a mind-boggling management task. In today’s serials environment, old and familiar problems remain, and they are accompanied by an entirely new set of complexities.

In his new textbook, Thomas Nisonger covers all of the above issues and more. Management of Serials in Libraries is a well-researched, comprehensive, up-to-date look at serials management issues. The emphasis is on collection management. This book appears on the heels of several other publications that also address the uniqueness of serials and the special handling that they require. Two recent notable works are Serials Management: A Practical Guide by Chiou-sen Dora Chen (1995) and Marcia Tuttle’s Managing Serials, with chapters by Luke Swindler and Frieda B. Rosenberg (1996). Chen’s publication, as its title suggests, is a concise guide of 186 pages that provides basic, straightforward information for efficient serials management. In contrast, Nisonger’s book is not a guide, but rather an in-depth study of serials with an intent “to educate rather than train” (xxi). As the author himself suggests, his book nicely complements Tuttle’s work by providing separate, detailed chapters on electronic journals, serials automation, citation analysis, and collection management issues.

Nisonger has an admitted bias toward academic libraries because they reflect his background and concerns as an associate professor in the School of Library and Information Science at Indiana University; however, the book is also intended for use by serials managers in public, school, and special libraries. It is written from “the perspective of a library and information science educator rather than a library practitioner” (xxi). Throughout his book, Nisonger consciously and effectively strives to create a balance in describing how serials are actually managed.
versus prescribing how they ought to be managed.

The book consists of ten chapters. Brief annotations that introduce each chapter and bold sub-topics within chapters allow easy browsing. Many of the chapters (and some of the sub-topics) end with a convenient summary and/or conclusion. An unannotated bibliography complemented by a list of further reading follows each chapter, and at the end of the book, the author provides a bibliography on serials in general. Immediately following the introduction, common abbreviations and acronyms are spelled out. This is a helpful feature, though there is no formal glossary that would be handy for quick and easy reference to serials jargon. Nisonger notes that this omission is intentional because terms needing explanation are defined as they appear in the text. The three useful appendixes are annotated bibliographies of sources for serials statistics and of pertinent serials bibliographies, and a list of Web sites relevant to serials.

As is traditional with most books about serials, the first chapter is devoted to definitions and descriptions of different types of serials, including electronic journals. Chapter 2 is an interesting account of the historical development of both print and electronic serials. In this chapter, the author highlights the impact of serial costs on library budgets. Nisonger's personal interest in collection management issues is revealed by his extensive coverage of the subject in two distinct chapters. These chapters highlight such topics as selection, deselection, document delivery, holdings gaps, treatment of unbound issues, and multiple copy and location decisions. He discusses microevaluation (a title-by-title approach to the make-up of the collection), and macroevaluation (a look at the collection in its entirety). Brief arguments relating to the concept of access versus ownership are presented with Nisonger's conclusion that, in the context of collection management, a desirable solution would be an appropriate balance between the two.

Nisonger prefers to use the term collection management rather than collection development because the word management can be more closely associated with the control of funds allocated to the serials collection. According to the author, budgetary concerns are more than ever a major factor in collection management decisions. Whereas a serials manager at one time needed only to decide whether or not to subscribe to a specific title, the serials manager of today must also take into consideration options for print-only subscriptions, electronic-only subscriptions, combined print and electronic packages, and document delivery services. Nisonger states that “interest in serials collection management has increased in recent years because of continuously escalating serials cost in tandem with stagnant library budgets, the emergence of new electronic formats, and an increased emphasis on access rather than ownership” (53).

In logical order, chapters on citation analysis, periodical use, and journal ranking studies follow. Nisonger defines citation analysis as a component of bibliometrics, “the application of quantitative techniques to bibliographical units, such as books or serial articles” (121). For use in collection management, citation analysis assumes that a cited item must have been used as a reference source in research for an article. Examination of citation patterns helps librarians meet the research needs of their patrons. The author explains in great detail the use of Journal Citation Reports, published by the Institute of Scientific Information, as it relates to serials collection management. He also describes Bradford’s Law, a pattern in the use and citation of journal articles, and its application to collection management issues.

Unlike other recent works on serials, Nisonger devotes an entire chapter to periodical use studies. He delves into the purposes of such studies and into various methods for carrying them out. He then describes how to tabulate the final results (cost per use) of the study and concludes with some facts about major periodical use studies that have been conducted over the years. Because of general interest in this topic, questions about different aspects of periodical use studies appear frequently on electronic discussion lists.
Any serials manager considering a use study will find this section of Nisonger's book an invaluable aid in planning and implementation.

Fundamental information about serials processing, electronic journals, and serials automation make up the remainder of the book. In the epilogue, Nisonger explores trends that may cause familiar issues of serials management to change even more in the future. He raises the question of the future of libraries in the context of the much broader question—what defines a library?—and then proceeds to question the future of the serial in its traditional format. Nisonger presents thought-provoking issues that deserve consideration by any dedicated serialist.

It is obvious in the vast number of notes and references that this book is based on complete and careful research in the customary sources and also in non-traditional sources, including electronic discussion lists, electronic journals, and the Web. The result is a well-organized source of information about the past, present, and future impact of serials on libraries. But first and foremost, Nisonger provides an in-depth look at current developments and available options in a rapidly changing environment for overall management of serials. This book is a worthy addition to the book shelves of "library practitioners whose professional duties involve serials...as well as library and information science teachers and researchers" (xviii).—Sylvia O. Martin (martin@library.vanderbilt.edu), Vanderbilt University Library, Nashville, TN


This collection of conference papers is an excellent resource. Aptly titled "visualizing" subject access, the collection is not an attempt at a final word on the subject, but is instead a set of works in progress, case studies, current experiments, and theoretical analyses of how subject content can be "seen." Included are twelve short papers, three abstracts (two with more complete Web sites), and opening and closing summaries. While this work does not satisfy the unrealistic desire for a single, complete solution to problems of subject access, the individual authors hold out the hope that many provisional solutions and continued creative experimentation will allow us to make significant progress. The authors are both older and younger scholars in the field and a cross section of information theoreticians, computer interface specialists, librarians, library school faculty, and vendors. While it is not possible to achieve in print the synergy of the conference milieu or the range of demonstration formats included in the conference presentations, the editors partly bridge this gap by providing Web addresses that extend some of the presentations, though they fall short of the yet unrealized multimedia electronic book in which real audio and video might work this magic. Here are some highlights from the print version of the conference.

A key paper is "Information Analysis in the Net: The Interspace of the Twenty-First Century," in which Bruce Schatz boldly predicts that within ten years "people will be able to solve real information problems themselves" (111), correlating information and doing analysis rather than merely searching. His vision is a system that draws on smaller and larger repositories, automatic as well as human indexing, interactive vocabulary switching, and "peer-peer not client-server" communications (123). Schatz uses the term "telesophy" to express a vision of the future knowledge community, growing and connected by an integrated conduit where the switching and technology are seamless and invisible to the average user, as in the current telephone system. Drawing on the history of development over the last ten years, his own work, and current experiments in progress, Schatz makes his predictions seem reasonable rather than far-fetched.
science fiction. His analysis provides a backdrop for the other papers. In turn, the other authors bolster some of Schatz’s speculation.

Elizabeth Liddy provides insight into the theory of natural language processing (NLP) in an extremely interesting paper titled “Natural Language Processing for Information Retrieval and Knowledge Discovery.” She articulates six linguistic levels (morphological, lexical, syntactic, semantic, discourse, and pragmatic) that must be analyzed in true NLP. Her exposition of the levels conforms with linguistic theory, and her descriptions of the capabilities of her Dr-Link, CHESS, and Know-it systems—she calls them “knowledge products”—is fascinating, but the implication that these products are capable of automatic analysis at all linguistic levels remains in question. At the conference she admitted some use of human indexers. Other contributions to theory include a paper by Raya Fidel and Michael Crandal on “The Role of Subject Access in Information Filtering.” This case study shows that users choose journal articles by a process that involves additional parameters beyond, yet related to, subject. The authors suggest that these additional parameters could be designed into a system. Bryce Allen, in “Visualization and Cognitive Abilities,” also makes a key point: any system for visualization of information must be designed in consideration of the variety of cognitive abilities and user preferences.

Several of the authors address vocabulary issues. In “Dimensions of Discriminability: The Role of Controlled Vocabulary in Visualizing Document Associations,” David Dubin describes the VIBE system that maps clusters of related concepts to ever-finer sets of documents. Two articles that should be read side by side, “Thesauri in a Full Text World” by Jessica L. Milstead and “Building and Accessing Vocabulary Resources for Networked Resource Discovery and Navigation” by Joseph A. Busch, provide a theoretical look at the future use of thesauri and a model system used at the Getty Institute for mapping across interrelated vocabulary structures.

Nicholas Bellin provides insight into particular systems for information retrieval and access system design in “An Overview from Rutgers Investigations of Interactive Information Retrieval,” with an important discussion of the Text Retrieval Conference programs and their evolution in the 1990s. Another avenue into experimental systems is Gerry McKiernan’s “The Big Picture (sm): Virtual Browsing in Web and Non-Web Databases.” McKiernan’s contribution to this book is not a full paper. Its abstract provides the URL of his more complete list of experimental projects (at http://www.public.iastate.edu/~CYBERSTACKS/DCP97.htm).

A final exceptional highlight of the book is Eric Johnson’s brief description of the IODynie interface. This Internet client program provides all of the desirable features of any information retrieval system, including querying multiple databases across language and protocol, common searches in the same form in multiple databases, persistent searches cached for later re-use, drag-and-drop navigation of multiple query spaces, and an information object structure.

In sum, this is an excellent work to consult for current thinking about subject access, about how users approach information, about how best to present information, and about what the resource landscape may look like in the next few years. It compares favorably to the Association for Computing Machinery, Special Interest Group on Information Retrieval conference proceedings (ACM SIGIR), but comes without the technical formulas common in SIGIR publications. Still, there is useful information for the more technically oriented information retrieval specialist. Papers are short, well indexed, and can be easily understood. Throughout the book, the authors use clear language, illustrate with creative diagramming, and provide follow-up references. —William J. Wheeler (william.wheeler@yale.edu), Yale University Library, New Haven, CT
In Memoriam: Phyllis Allen Richmond

Nancy Williamson

On October 6, 1997, Phyllis Allen Richmond, internationally known librarian, information scientist, and educator, died from complications of Alzheimer's disease. She was 76 years old. At the culmination of her career, she was a professor emeritus of Case Western Reserve University in Cleveland, Ohio, where she served from 1970 to 1984 as a distinguished expert in classification theory and practice.

Born in Boston and raised in Rochester, New York, Phyllis Richmond has an educational profile with more than fifteen years of distinguished scholarship. In 1942 she received an A.B. in history from Western Reserve University, where she was elected to Phi Beta Kappa. Following a semester as a graduate scholar at Bryn Mawr College, she served as curator of history at the Rochester Museum of Arts and Sciences from 1943 to 1945 and again in 1946/47. Continuing her education, in 1946 she received an A.M. degree in history from the University of Pennsylvania. In 1947/48 she was an American Council of Learned Societies fellow at Cornell University, and in 1948/49 was a Bennett fellow at the University of Pennsylvania, where she received a Ph.D. in the history of science. Following this, she was a research assistant to the director of the Institute of the History of Medicine, Johns Hopkins University. Then in 1952 she entered the School of Library Science at Western Reserve University, receiving an M.S.L.S. degree in 1956, and was elected to Beta Phi Mu. Dr. Richmond's academic excellence, together with her intense interest in history, science, medicine, and information science, was to have a profound influence on her career as reflected in her professional activities, her research, and her writings.

Phyllis Richmond started her career as a librarian at the University of Rochester, where she began as a serials cataloger (1955 to 1960), then served as supervisor of the River Campus Science Libraries (1961 to 1966) and as information systems specialist (1966 to 1968). Innovation was a factor early in her career, when she had a strong interest in the application of computer technology. While at Rochester she was responsible for six computer-produced title-a-line book catalogs (1960 to 1968) for the science and engineering libraries and for two computer-produced serials lists (1966 to 1968), all of which were leading edge projects for that time. Following a brief period in 1966, when she was a visiting professor at the School of Library Science, Western Reserve University, Dr. Richmond moved from a career in librarianship to library and information science education, first at Syracuse University in 1969 and then at Case Western Reserve University from 1970 until her retirement in 1984.

Dr. Richmond's career as an educator was marked by the fact that she was an outstanding teacher of cataloging and classification and guided numerous Ph.D. candidates through their programs. In particular, she brought to her teaching a

Nancy Williamson (william@fsu.utoronto.ca) is professor emerita, Graduate Studies in Information Studies, University of Toronto.
rare mixture of theory and practice, and she was arguably the most outstanding classification theorist in North America for her time. She had the ability to present complex topics with clarity and wit and to instill in her students an enthusiasm for their work heretofore unknown to them. In her role as an educator, she endeavored to cultivate in her students an understanding of classification and classificatory structure as the basis for the organization of all knowledge and indeed all aspects of life in general. She was a firm believer in the fact that “some kind of classification is inherent in everything we do” (Richmond 1977, 108). An early supporter of interdisciplinary study, she encouraged students to think, to analyze and compare, and to seek logical solutions to problems. For teaching purposes, she frequently drew analogies between library classification and the rest of the world, inviting comparisons between scientific subjects and cataloging and classification. For example, she once posed the following problem for one of her Ph.D. students: “There are many kinds of medicine based on different theoretical foundations: regular, homeopathic, allopathic, osteopathic, naturopathic, chiropractic, etc. All practicing medicine uses the same materia medica (i.e., if licensed by a state they can write prescriptions for drugs). What are the equivalents in classification and cataloging?”

As a library and information science educator, Phyllis Richmond also took an active part in professional associations, most particularly in ALA and the American Society of Information Science (ASIS). In ALA, she was an inveterate participant in ALCTS, bringing to them her interest in practice and her knowledge and expertise in research and library and information science education. She served as chair of the CCS Classification Committee, member of the CCS Policy and Research Committee, member of the CCS Executive Committee, secretary of the Serials Section, and chair of the Esther J. Piercy Award Committee. For many years, among her many commitments was her work as assistant editor of LRTS, where she refereed many papers and encouraged many aspiring researchers. She was also very generous with her time and effort in contributions to ALCTS programs, workshops, and classification institutes. An active member of ASIS, Dr. Richmond was one of the American pioneers in documentation and information science, serving as the first chair of its Classification Research Special Interest Group (ASIS SIG/CR). Indeed her expertise and reputation took her beyond these two bodies to serve as member of two subcommittees of the Committee Z39 of the American National Standards Institute, as consultant on various information science projects at the American Institute of Physics, and on the Bibliographic Control of Microforms Project of the Association of Research Libraries.

Dr. Richmond was also highly regarded internationally in the field of classification theory. She was a corresponding member of the British Classification Research Group (CRC), which strongly influenced her involvement in the formation of a Classification Study Research Group (CSRG) in the United States. She also presented papers at major international conferences, among which were the FID/CR International Study Conferences on Classification Research. In this context, she participated at the second study conference in Elsinore, Denmark, in 1961 and prepared a paper for the third conference, which was presented in Bombay, India, in 1975. Although the CSRG did not have a long life, its aims and objectives are somewhat embodied in the activities of the ASIS SIG/CR and may be regarded as an important part of Phyllis Richmond’s legacy to classification research.

With respect to her writings, her multifaceted expertise and creative mind made it possible for Dr. Richmond to gain deep insights into the theoretical aspects of knowledge and to arrive at new and intriguing methods of viewing her subject. She was prolific and truly interdisciplinary. By the time she retired, she had produced more than 75 articles and a book. She always had fresh approaches to her subject, and a number of her writings and presentations are unique in this regard.
For example, her article on “Cats: An Example of Concealed Classification in Subject Headings” (Richmond 1959) reflected not only her pleasure in her beloved cats, Fluffy and Brownie, but also it brought together her knowledge of zoology and the problems of structure in the Library of Congress Subject Headings. Similarly her keynote address, titled “Mr. Dewey's Classification, Mr. Cutter's Catalog and Dr. Hitchcock's Chickens” presented at the 1976 ALA/RTSD meeting as a part of the centennial of ALA, Dewey Decimal Classification, and Cutter's Rules for a Dictionary Catalog, provided much food for thought for her audience. The paper drew on the study of dinosaurs to illustrate the importance of research as “an activity that constantly reinterprets the past as well as the present” (Richmond 1977, 108) in the development of the principles that should underlie the work that catalogers and classifiers do in practice. It is an article that, more than twenty years later, still has much to say to library and information science. Always looking to the future in her research and teaching, Dr. Richmond was a staunch supporter of modern methods of classification. In particular, she had become the North American expert on faceted classification and the work of S. R. Ranganathan and the CRG. As a basis for the study of classification she produced a "Reading List in Classification Theory" (Richmond 1972). This bibliography was structured in such a way as to provide a well-organized and useful approach to a thorough study of this topic. Ultimately her continuing interest in modern classification systems led her to an investigation of the PRECIS system, a product of the work of the CRG, and resulted in her very lucid and useful Introduction to PRECIS for North American Usage.

For her energy, her scholarship, and her intellect, Phyllis Richmond received numerous awards of distinction that were justly deserved. In addition to her scholastic awards, she was honored professionally. In 1968, the American Documentation Institute (predecessor of ASIS) awarded her its Technical Service Award, and in 1972 she was the first woman to receive the ASIS Award of Merit for her contribution to the understanding of the theory and practice of subject analysis, in general, and classification in particular. Then, in 1977, she was honored with ALA/RTSD/CCS’s highest award—the Margaret Mann citation as an outstanding individual in the fields of cataloging and classification, as a writer, scholar, practitioner, teacher and contributor of common sense and wisdom.

Indeed, during her lifetime, Dr. Phyllis Allen Richmond made an outstanding contribution to the development of theory and practice of classification—a contribution on which her successors have been able to build and expand. She was exceptional for her time and has left a lasting legacy that should not be forgotten. In all of this, Phyllis was a quiet unassuming person who enjoyed life through her many friends, colleagues and students, her beloved cats, and her ham radio. She is greatly missed both personally and professionally. Hers was an act that will not be easy to follow.

REFERENCES
———. 1972. Reading list in classification theory. Library resources & technical services, 16: 364-82.
———. 1977. Mr. Dewey's classification, Mr. Cutter's catalog and Dr. Hitchcock's chickens. Library resources & technical services, 21, 107-19.
BUILD THE ASIAN LANGUAGE COLLECTIONS YOUR USERS WANT WITH THE OCLC ASIALINK AND TECHPRO SERVICES

Developing Asian language collections is easy with the OCLC AsiaLink service. Staffed by Asian language experts, AsiaLink offers everything you need to establish, maintain and expand Chinese, Japanese, Korean, and Vietnamese (CJKV) collections in your library.

- The AsiaLink Collection Development service lets you determine the specific subject matter and format of materials that AsiaLink staff will acquire on your behalf, based on your users' needs. You'll also get significant savings on the cost of books, including higher-than-average discounts from book vendors.

- AsiaLink Sets get you off to a quick start or allow you to easily expand existing collections. Books are available in sets of 25, 50, or 100 pre-selected fiction and nonfiction titles. Both juvenile and adult sets are available. Title lists are updated annually, so it's easy to add to your collection regularly.

- Full-level OCLC cataloging is provided for all titles and materials may be physically processed upon request.

For help cataloging titles not acquired through AsiaLink, turn to the OCLC TechPro service. TechPro provides expert cataloging for materials in all bibliographic formats and in many languages (including CJKV), plus optional physical processing. TechPro can help you quickly and efficiently eliminate backlogs and keep your cataloging up-to-date, giving your users prompt access to materials.

For more information on AsiaLink or TechPro, contact Marcia Stout at OCLC.

1-800-848-5878, ext. 4386  614-764-4386  Fax: 614-718-7277
E-mail: marcia_stout@oclc.org

www.oclc.org/oclc/menu/col.htm