The Need for Funded Research

Tschera Harkness Connell *an assistant professor for the* school *of Library and Information Science, Columbus Program at Kent State University.*

Part I*: Barriers to Effective Subject Access in Library Catalogs*

The first part of this article is a brief summary of an OCLC-funded project, "Identifying Barriers to Effective Subject Access in Library Catalogs," in which I participated. The project was under the leadership of Professor F.W. Lancaster of the University of Illinois at Urbana-Champaign. A more complete report of the project is published elsewhere (Lancaster, Connell, Bishop, and McCowan, 1991).

Purpose

The purpose of the study was to determine the probability that a skilled catalog user would retrieve "the best" materials represented in the catalog on some subject and, if they are unable to retrieve the best materials, to determine what changes would be needed to ensure that future catalogs would allow the user to retrieve more of the better materials.

Background

In most studies on how to improve subject searching in online catalogs, success is measured in terms of whether the user is able to match subject terminology with the terminology of the catalog, or whether or not the user selects an item or items from among the items retrieved from the match. Such definitions of success do not consider whether or not the user has located *anything useful.* They do not address the issue of whether the user can locate what is in some sense best (i.e., the most complete, the most up-to-date, or the most authoritative).

Traditionally, the purpose of the catalog is not prescriptive. In fact, one of the purposes of the catalog is to present all the related works in the collection. However, it is the assumption of this study that users want to be able to locate what in some sense is the best. "Best" in this study is defined as "recommended." We examined whether or not skilled users would retrieve books that appeared on lists, compiled by specialists, of recommended readings in various subject areas.

Methodology

Fifty-one bibliographies on a wide range of topics were assembled. The lists were obtained from faculty and from recommended readings appearing in recently published articles in encyclopedias or encyclopedic dictionaries. The sample of topics used in the study was determined by the availability of faculty lists and fairly recent (1983-1989), specialized bibliographies containing significant numbers of items likely to appear in the catalog of a research library.

For each bibliography, the following steps were taken:

1.   Journal articles were eliminated, since traditionally these have not appeared in library catalogs.
2.   A search on the topic of each bibliography was performed in the "full" online catalog of the University of Illinois (FBR) by two members of the team who were familiar with FBR. These two members performed all the searches. The searches were performed on the basis of the title of the article or bibliography only. The searchers did not see the bibliography until after the search was completed.
3.   Items in the bibliography not retrieved by the subject search, and then subsequently determined not to be owned by the University of Illinois, were eliminated.
4.   Items not retrieved by the original subject search were gathered and examined to determine why items presumed to be relevant to a particular topic were not retrieved in the original subject search. An evaluation was made to determine how the search strategy or characteristics of the catalog would have to be changed to allow these to be retrieved.

Results

The results of the fifty-one searches varied from eight cases having 100 percent recall to two searches with zero recall. The fifty-one bibliographies collectively contained 607 items included within FBR, and of these, 327 were retrieved in the subject searches. If we simply average these numbers (327/607), we get an average recall of 53.9 percent.

This result is probably higher than what most users would achieve. The searchers were instructed to search broadly, which means that the searchers used all seemingly relevant terms that they identified in LCSH, at any level of specificity, and not just the broadest applicable subject headings. Thus, a search on Pre-Columbian religions included terms related to specific religions as well as the more general terms.

The searchers were further instructed to give no concern for the precision of the search. For example, to get a high recall on the Gumbel distribution, which relates to the statistics of extremes, the searchers used broad terms such as *Mathematical statistics* and *Stochastic processes,* which retrieve records for more than 1,200 items. This same situation applies to other searches. Therefore, while recall was high in a few of the fifty-one searches, these results would not be achieved under real-life conditions because a library user would just not be willing to look through hundreds of records to find a handful of items.

High recall and high precision occurred only in situations where the subject of the search coincided closely with a subject heading or headings. For example, the search on the image of women in the Bible achieved 75 percent recall on the single heading *Women in the Bible* and could have achieved 100 percent recall by the use of the additional term *Woman (Theology).* Such close matches between subject heading and topic were rare.

The main purpose of the study was to determine what might be done to library catalogs to make them more effective tools for subject access. With this in mind, items that were not retrieved by the initial subject search were examined in order to determine what changes in the catalog and/or indexing policy would be necessary to make it possible to retrieve the recommended items (see Table 1).

How Results Could Be Improved for the Fifty-one Searches

Table I.

Total number of relevant items in FBR for fifty-one searches	607	
Number of relevant items retrieved in fifty-one searches	327	
Average recall ratio (327/607)	53.9%	

Possible Improvement	Additional Items Retrievable	Revised Recall
Elements in the bibliographic record		
Closely related and somewhat related subject headings	51	62.3
Other parts of record	10	55.5
Subtotal	61	63.9
Enhancements to record		
Indexes of books	125	74.5
Contents pages	86	68.0
Full text	58	63.4
Subtotal	211*	90.3
Not retrievable even on full text	8	

*Categories "Indexes of books" and "Contents pages" are not mutually exclusive.

In looking at the subject headings assigned to the items, it was determined that had we searched on closely related headings, the recall ratio would have increased to 378/607 (62.3 percent), an improvement of a little more than 8 percentage points. For example, had we searched on the heading *Glossolalia* for the topic "spirit possession" or the heading *Poverty—Government policy—United States* for the topic "hunger and malnutrition in the U.S.," recall would have slightly improved. Of course, this is theoretical improvement based entirely on hindsight.

If the searches had been broadened to include titles and other information, little improvement in recall would have occurred. Only ten of the 229 items not retrieved by subject headings would have been retrieved. That extending a search from subject headings to titles or

subtitles has minimal effect on recall suggests that the subject headings assigned are very "close" to the terminology of the titles.

The results shown in Table 1 might suggest that the problems of subject access in library catalogs could largely be solved if the text of contents pages and/or indexes were stored in a form suitable for searching. Nothing could be farther from the truth. Many searches on extended records would retrieve thousands of items rather than the hundreds that were retrieved in many of the searches on existing records alone. Only in the case of an atypically specific search, involving a rather rare word or name (such as Gumbel), might the enhanced record improve search results.

The library catalog, as it now exists, may provide adequate subject access for a small collection, but it is inadequate for a large, multidisciplinary library

Implications

In this study comprehensive searches were performed in order to determine to what extent the items considered important by the expert could be retrieved by the persistent and diligent searcher. Some of the failures were due to factors other than indexing policy and catalog design. Examination of recommended items showed that a significant number of the readings recommended by experts are relevant by analogy only.

Although these items may be important to the topic, it is difficult to see how they could have been retrieved by any likely search approach. For this reason alone, if one wants to know the best things to read on some topic, there is no substitute for consulting an expert, either directly or indirectly (e.g., through an expert-compiled bibliography). What is discouraging about the results of this study is that items clearly relevant to the topic are often difficult to retrieve.

Library catalogs as now designed permit only superficial subject searches. Lack of specificity in subject headings, coupled with the fact that the catalog provides access only at the level of the complete bibliographic item makes it virtually impossible to achieve a high recall at an acceptable level of precision. Ability to search on other parts of the bibliographic record (e.g., keywords in title), and even to access a greatly enhanced record (e.g., containing contents pages or book indexes) does not improve performance as much as one might expect. The library catalog, as it now exists, may provide adequate subject access for a small collection, but it is inadequate for a large, multidisciplinary library.

Despite popular belief, the transformation of the card catalog into an online database has not significantly improved subject access. Indeed, it may have made the situation worse because it has led to the creation of much larger catalogs that represent the holdings of many libraries. The present study shows solutions must be sought in providing ways for the user to browse large groups of retrieved bibliographic records, to discriminate among the records retrieved, and ultimately to choose from among the items themselves.

Partial solutions to the problems may involve the adoption of detailed analytical subject cataloging, and the development of finer tuned vocabulary control. Users need ways to explore categories of related headings without having explicitly to think-up or key-in all relationships. We must continue our research efforts to explore both conceptual and mechanical solutions to the difficulties of achieving both high recall and precision in the online catalog.

Part 2: *An Educator's View of Research*

The second part of this article gives one educator's view of research. This view is described in the context of the education and research roles of the library and information science educator.

The University Research Environment

Although research can be triggered under a fruit tree with little more than an inquiring mind and a falling apple, promoting research requires creating an environment that nurtures ideas and encourages participation. Ideas for research abound[1] and collegial support within the university provides the natural forum for the discussion and nurturing of ideas.

Participation is encouraged with the provision of the tools and resources that enable faculty to accomplish research. Adjusting faculty workloads through the use of research assistants, laboratory monitors, and adjunct faculty, plus minimizing committee work, are some of the ways that administrators have traditionally encouraged faculty participation in research. The building of library research collections has been one way of providing tools.

However, universities typically do not do as well in the provision of scientific equipment, computers, or software. Because of limited funding, universities often are unable to keep up with industry in terms of state-of-the-art technology. While it is possible for industry to pass the cost of upgrades on to customers, academic departments within universities are more limited in their funding options. This cost limitation puts university researchers at a disadvantage. Inadequate funding may limit the questions that can be asked, or at least the ways in which the questions can be examined. Inadequate funding also limits the training that can be passed on to students.

In the fields of library and information science, technology is critical. Our research agendas call for examination into issues such as access to information, the information needs of users, the uses of information (U.S. Department of Education. Office of Educational Research and Improvement, 1988, pp. 16-19), the impact of technology on research, and preservation studies (Lynch and Young, 1990, pp. 235-240). All of these require interdisciplinary approaches (for which the university provides an excellent setting) and state-of-the-art technology (for which the university is hard pressed to keep up). As university budgets continue to shrink, the funds to pursue these questions will need to come from outside the university.

Teaching Students Research Skills

Performing research is just one aspect of the educator's role; teaching students is another. For many of us, it may well be that our most enduring contribution will be the impact that we make on our students. In this context of teaching, let me describe what I consider to be a somewhat schizophrenic role of the library and information science *educator*.

Most of our students are masters students, at the beginning of their professional careers. As educators, it is crucial that we enable these students to be intelligent consumers of research. We must help students learn to evaluate research well enough to be able to incorporate useful findings into their thinking and into their decision making. It is even *more* important that we help our students develop an enthusiasm for research. Basic research skills will aid them in problem solving, marketing, and planning. However, inculcating enthusiasm is very difficult to do. Certainly it helps if we are enthusiastic about our own research. But the reality is that for many students the immediate concern is how to get a job and how to perform on that job.

I teach in the area of technical services. In cataloging courses, the student preoccupation with the immediate translates to more of an interest in "how to" than "why." Yet, if we do not address the broad issues of information management, our students are not prepared to evaluate new approaches to organizing information for access. If we do not help students understand why particular solutions have been adopted in the past, they are less likely to understand the complexities of the issues that new technologies present. Finding the proper balance between

theory and practice is often difficult. The ultimate goal must be to teach students the skills and attitudes needed for self-directed inquiry. Teaching the skills and attitudes for self-direction is the means for unifying the theory/practice schizophrenia. These are the skills and attitudes that are essential to a competent practicing professional information specialist. These are *also* the skills and attitudes that are essential to persons involved in research.

The tension between practice and theory is usually not a problem with doctoral students. The chances are good that our doctoral students already have a commitment to research by the time they enter the Ph.D. program. Many are established in the field before they begin the degree. Because of this, their needs are similar to those of faculty, but just a step removed. They need nurturing, time, and tools. In addition, they need opportunities to work with experienced faculty engaged in research.

Obviously, one of the ways to get students (at any level) excited about research is to get them involved in the process—to provide for them a positive research experience. This can be partially accomplished through assignments and research papers. It can also be accomplished, at least on a limited scale, by involving students in faculty research projects.

The OCLC-funded project that I participated in involved a senior faculty member and three students. Two of the students were master's degree students working as graduate assistants. I was a doctoral student, participating out of interest. I found the interaction of the group process invigorating. It seems to me that the process has a number of advantages.

For the student, the group process provides a protective environment yet at the same time an intellectually challenging one. For the advancement of certain kinds of knowledge, the group process may be necessary due to the complexity of the issues under examination. The interdisciplinary nature of many of the questions facing library and information science researchers require a variety of skills not likely to be present in a single person.

_____

Much of OCLC's research has led to products of services that have greatly altered the ways that libraries provide their services._____

OCLC

Since 1986 OCLC, through its Library School Research Equipment Support Program and later its successor the Library and Information Science Research Grant Program, has funded approximately thirty such projects. The grant program "assists schools of library and information science to conduct high-quality technical research" by funding release time from teaching for the principal investigator, research assistants, travel, equipment, and other project-related expenditures" (OCLC, 1988, p. 41).

Applicants for the program must explain the significance of the research, why the research is innovative, and suggest future directions the research might take. From an examination of the variety of projects (both basic and applied) that have been funded over the years, it is apparent that this grant opportunity is a pretty open invitation—that projects need not be tied to OCLC products, services, or research agendas. OCLC is to be commended for that. Open-ended funding opportunities are rare.

We need more such funding. We need funding that can be used for basic research. For too long we have depended upon tradition within the field and commercial vendors outside the field to define our services. To take an example from recent history, during the early stages of the online catalog we were experimenting on our users with very rudimentary and inefficient automated catalogs. Since then, researchers have begun to look at the organization and retrieval of knowledge

more broadly in terms of how knowledge is created, stored, and used.

However, more exploration is needed on how people seek information. More testing is needed to see how proposed solutions work in different environments, (e.g., disciplines, institutions, cultures). It is the basic research that will give us the depth of understanding needed for future, viable applied research. It will be a strong foundation of basic research that will eventually enable us to design and build the tools that provide access to ideas— whether the ideas are accessed through people, through print, or through processors.

Over the years, OCLC has been heavily involved in research. OCLC itself is a product of research. For example, cataloging workflow studies have led to products that have streamlined cataloging operations and have changed the day-to-day operations in technical services.

More recently, OCLC research into using the online union catalog for collection evaluation has led to a new service, the OCLC/ Amigos collection analysis system. However, OCLC has also been heavily involved in basic research. Research into retrieval techniques, document structures, and interface design and management may (and probably will) result in products, but for the present, research in these areas increases our basic understanding of the interactive effects of language and text, and of presentation and retrieval techniques.

In the 1991 summary of OCLC's strategic plan, OCLC states that its strategy "is to return to the basics, to fundamentals"—to emphasize librarianship, collaboration and cooperation, to build upon its strengths in cataloging, resource sharing, and excellent user service (OCLC, 1991, p. 12). Library and information science faculty have similar interests in building foundations. I would encourage library and information science educators and OCLC to look for ways of increasing opportunities for OCLC/faculty collaboration.

Perhaps schools of library and information science and OCLC could co-sponsor summer research institutes for groups of researchers to explore predefined topics of mutual interest. These institutes could be designed to promote interdisciplinary research.

They could be designed to promote practitioner/educator collaboration. If held at OCLC, these institutes could offer technical support not possible at most schools of library and information science. Such institutes would also provide a means for technology transfer back to the universities. Without the opportunity to use and evaluate state-of-the-art technology, it *is* difficult for faculty to know what is needed or to know for which to ask.

Another idea for cooperation would be to use schools of library and information science as test sites for developing OCLC products. This would be a way of increasing the research involvement of new recruits to the field.

There are many ways to take advantage of our shared research interests. I would encourage library and information science educators and OCLC to continue to explore ways of doing so.

Note
1. For examples in library and information science see "Research Questions of Interest to ARL" in U.S. Department of Education. Office of Educational Research and Improvement, 1988, *Rethinking the Library in the Information Age,* Vol. 1; "Research Questions of Interest to CLR" in Lynch and Young, eds., 1990, *Academic Libraries: Research Perspectives;* and "ACRL Research Agenda" in *College and Research Libraries News,* 51 (April 1990): 317-319.

References
ACRL Research Agenda. (1990). *College & Research Libraries News,* 51,317-319.

Lancaster, F.W., Connell, T.H., Bishop, N., & McCowan, S. (1991). "Identifying Barriers to Effective Subject Access in Library Catalogs." *Library Resources & Technical Services,* 35,377-392.

Lynch, M. J. & Young, A., Eds. (1990). *Academic Libraries: Research Perspectives.* Chicago: American Library Association.

OCLC. (1988). *Annual Review of OCLC Research, July 1987-June 1988.* Dublin, Ohio: OCLC Online Computer Library Center.

OCLC. (1991). *Journey to the 21st Century: A Summary of OCLCs Strategic Plan.* Dublin, Ohio: OCLC Online Computer Library Center.

U.S. Department of Education. Office of Educational Research and Improvement. (1988). *Rethinking the Library in the Information Age* (Vol. 1). Washington, DC: U.S. Government Printing Office.