Use of Circulation Statistics and Interlibrary Loan Data in Collection Management

Jennifer E. Knievel, Heather Wicht, and Lynn Silipigni Connaway

The authors analyzed the holdings, circulations, and interlibrary loan (ILL) borrowing requests of the English-language monograph collection at the University of Colorado at Boulder. Data for each area were mapped to conspectus subject areas, using Library of Congress Classifications, and then compared. The resulting data and subject distributions were analyzed by overall holdings, transactions per item, percentage of collection circulated, and a ratio of ILL holdings to requests. The method of analysis used in this study could be fruitfully applied to other research collections to assist with remote storage, preservation, and collection development decisions.

In the current economic environment, it is critical for librarians to practice evidence-based decision making. With decreasing library budgets, especially for monographic materials, it also is essential for librarians to assess and manage collections to determine subject areas for acquisition, de-accession, digitization, preservation, and remote storage. One empirical collection assessment method is to examine usage statistics, such as circulation and interlibrary loan (ILL) data. Many librarians employ usage data as an indicator of the library's materials' relevance to user populations.

Sonia Bodi and Katie Maier-O'Shea believe that librarians' knowledge of collections is often intuitive and urge librarians to make data-supported collection decisions. They suggest that collection development focus on learning outcomes of library users rather than on strictly comprehensive collections or even curriculum support.[1] This approach is impossible without solid data from various perspectives.

Librarians at the University of Colorado (UCB) at Boulder Libraries utilized circulation statistics and ILL data in the development of collection and assessment criteria that accurately reflect patrons' needs. UCB is a Research I, doctoral-granting institution with 26,000 FTE students. The libraries hold approximately three million volumes. Collection assessment has become more important with the current budget reductions at UCB, a problem shared by many libraries around the state and country. The materials budget has suffered severe permanent and one-time cuts that make efficiency in collection management increasingly important.

Like many university libraries, the UCB libraries' shelves have been filled to capacity for some time. In 1998, UCB began a remote storage project, and today, approximately 425,000 volumes are stored in a shared remote storage facility in Denver. The UCB subject bibliographers have used a variety of methods and tools to select materials to be placed in remote storage. One of these methods is the analysis of circulation statistics and ILL borrowing data.

OCLC Online Computer Library Center, Inc. (OCLC) maintains the WorldCat bibliographic database, the Online Union Catalog, which contains more than 50 million records. WorldCat serves not only as an aggregator of bibliographic data but also lends itself to data mining and data analysis. The WorldCat database includes the holding symbol for every member library holding each item represented in the database. There are almost a billion holding locations for library resources identified in WorldCat.

This study compares the UCB libraries' circulation statistics and ILL borrowing requests between January 1, 1998, and December 31, 2002, to UCB's WorldCat holdings data during this same time period. Four analyses are compared and discussed: the UCB's overall holdings; the average number of circulation transactions per item; the percentage of

the subject collections circulated; and the holdings to ILL ratio. All are compared using conspectus subject categories. These analyses can provide empirical data for collection management and remote storage decisions.

Literature Review

Reduced buying power in libraries makes the use of data in collection development of increasing importance. Difficult selection and funding decisions are more informed when they consider usage data. In 1995, Chuck Hamaker stated that collection for the sake of the collection is no longer possible except in the most financially insulated institutions.[2] Charles B. Osburn stated in 1992 that economic stresses and strains influenced a new set of guiding principles for library and information services.[3] One such principle is a shift in emphasis from spending acquisitions budgets on the basis of speculation to spending on the basis of demand. As Dennis P. Carrigan observed, this shift should increase the importance of data produced by automated systems.[4] Carrigan wrote that the recent crisis of scholarly publishing has made it necessary for librarians to prove that their libraries are making good use of the materials they have. The 80/20 rule (80 percent of users' needs are satisfied by 20 percent of the collection) is no longer satisfactory in this fiscal climate.[5] Carrigan believes that increased accountability in libraries will necessitate the use of circulation and other types of data.

A 1998 *Library Journal* survey found that "virtually every library LJ surveyed relied on faculty when making purchasing decisions, and half of them ranked faculty as the number-one source."[6] Although faculty suggestions should be considered, they should not be the only source of input. The library must satisfy the needs of all users. According to Hamaker, "both computer science and math faculty say don't buy books, we need journals; usage patterns say something quite different to the selectors."[7] He also speculated that

even in ARL libraries, some selectors still have the attitude that they know best what to buy and do not need to know what is being used (and who is using it). Hamaker has advocated the use of data in collection development, stating that intuition is not the key to understanding the usage of library collections.

Lynn Silipigni Connaway and Justin Littman state, "Circulation analysis is one of the traditional approaches taken to use studies and collection evaluation in libraries. The results of circulation analyses have been applied to a number of important issues, including evaluating collection acquisition policies, guiding such management decisions as allocating physical space for materials, identifying materials for offsite storage, allocating funding for materials, and suggesting approaches to deselection."[8]

Although the collection and analysis of circulation data have been fairly common practices for some time in libraries, an examination of the existing literature on the topic suggests that circulation data are not being widely used in combination with other types of data for collection development decision making. The articles that do discuss data-based collection development tend to consider only circulation data, not ILL data or a combination of both types of data. Mike Day and Don Revill stated that "circulation is the strongest single element we have on which to base decisions and we feel that the analysis of this data provides a useful tool; a tool that could be used by our subject librarians along with their professional judgment and user satisfaction surveys to evaluate and fine-tune their collections."[9] In an article describing a variety of methods of collection evaluation, George S. Bonn introduced a method of obtaining proportionate circulation statistics by subject class. The number of circulations in a given subject area is compared to the number of holdings in the same subject area, resulting in a ratio Bonn labeled the "use factor" for that subject area.[10]

William Aguilar wrote one of the first articles that advocated the combined use of circulation and ILL statistics in collection development.[11] It described a series of basic methods for comparing holdings data, circulation statistics, and ILL statistics to assess the activity of the library collection. He recommended analyzing collections by subject areas and described three primary methods for evaluating subject areas. Aguilar suggested that Bonn's use factor can be applied to determine the number of circulations relative to holdings. He further explained that Bonn's use factor takes on greater significance as refined by Terry R. Mills's "percentage of expected use," which is simply the use factor multiplied by one hundred to obtain a percentage.[12] Subjects above 100 percent can be defined as overused whereas subjects below 100 percent are underused. Aguilar introduced the concept of a similar ratio to analyze ILL borrowing requests. The "ratio of borrowings to holdings" compares the number of interlibrary loans relative to the holdings in a given subject area.

John N. Ochola has used the methods described in Aguilar's article to conduct a pilot project at Baylor University to determine which monographs should be moved to remote storage.[13] He also applied Aguilar's four "rules," which advocate specific actions for subjects with low or high ILL borrowings and low or high circulation.

Bonn, Aguilar, Hamaker and Carrigan all discussed the evaluation of circulation and ILL data by subject area. As Carrigan stated, "Circulation data must be analyzed at the macro level (across subject areas, not title by title) to bear significance on resource allocation and guide collection development."[14]

Methodology
Data
This project utilizes three distinct sets of data from the UCB libraries. In all three sets, data are limited to books only and analyzed across subject areas, rather than

title by title. It is important that all three data sets employ the same definition of a book so that overlap can be meaningful rather than misleading. For instance, if one set includes government documents as "books," but another does not, comparing the data based on subject, circulations, and requests will no longer be valid. (For the purposes of this project, a specific definition of "books" is discussed later in this article.) The three data sets are:

• The entire collection of books owned by UCB

• All the books that circulated at UCB from January 1998 through December 2002

• All the books that were requested via UCB's ILL department between January 1998 and December 2002

A number of items are excluded from the data sets. The UCB Law Library is a separate organization from the rest of the UCB libraries and therefore is not included in the study. The University of Colorado's other campuses in Denver and Colorado Springs also are not included. This study is limited to English-language books; foreign-language materials comprise nearly 25 percent of the total ILL requests, constituting a large enough data set for separate study. Government documents are excluded from this study because the extent of government documents collections is generally determined by depository status, rather than active collection development decisions, making comparisons of circulation, holdings, and borrowing requests irrelevant. Manuscripts, dissertations, and theses also are excluded from the data because many unpublished items are generally held by only one institution. In particular when considering ILL borrowing requests, it is misleading to include unique items such as dissertations in the source data to compare against the holdings of a library that would not reasonably be expected to acquire the item. Musical scores, although representing a respectable proportion of borrowing requests, also are excluded because they are not language material and

represent a different kind of user need and behavior than this study attempts to analyze. Wherever possible, OCLC numbers were used to obtain Library of Congress Classifications (LCCs) for all items, for the purpose of subject area designation. Items without OCLC numbers are classed as "unknown."

Definition of a Book

Though it seemed on early consideration that a "book" was an obvious entity, it became clear that this study needed a very specific definition of a book. The operational definition of a book used for this research was as follows: a book is an English-language printed monograph, including large print and print reproduction. Several criteria were used to exclude materials in any medium except paper:

1. Bibliographic Level is *m*, or *monograph*, in the MARC record and MARC 007 field is blank, *a, b*, or *d*. These requirements excluded nonbooks such as serials, government documents, dissertations, theses, manuscripts, scores, and music.

2. Items do not have a $h field in the 245 MARC field and blank, *d*, or *r*, in the 008 field. These requirements excluded items in any medium except paper, such as microfilm, and any nonbook materials except large print and print reproductions.

3. Items had a 260 $b or 020 field and a MARC 300 $a field with more than forty-nine pages. These requirements limited the data set to published materials of fifty pages or more only, and eliminated all theses, dissertations, manuscripts, pamphlets, and other unpublished works.

Conspectus Subject Categories

The data for this study were analyzed by subject classifications. The conspectus concept was developed by the Research Libraries Group, beginning in the 1980s. Conspectus subject categories are broad subject classifications based on the Library of Congress classification system that provide a basis for large research libraries to assess the subject distribution of

their collections. The conspectus classification has twenty-five subject areas. An existing research collection, in this case the collection at UCB, can be mapped to these conspectus categories in order to provide a picture of the overall concentrations in that collection. With only twenty-five subject categories, it is not sufficient to use the conspectus to evaluate a collection in detail. However, it is very useful to provide a broad picture of a collection. Table 1 lists all the conspectus categories. From this point forward, these subjects will be referred to as conspectus subject categories.[15]

Holdings Data

The authors pulled the UCB holdings data from holdings in OCLC WorldCat in July 2003. Again, the set was limited to books as defined above. Several elements of the UCB collection are missing from this data set, most notably, the holdings of the special collections and archives departments, most of which have not been cataloged. This does not present a significant obstacle to this study for several reasons. It is unlikely that holdings of special collections and archives represent a significant overlap with ILL borrowing requests because of the rarity of the items hold. They are noncirculating collections, so they could not be present in the circulation data. In addition, a dearth of electronic records in special collections and archives is a problem mirrored by nearly all major university collections for several reasons, including long-term backlogs in cataloging departments.[16] Therefore, the methods and results of this study, even without the special collections and archives records, still will be entirely relevant to other collections, indeed, even more so than if the two departments were included, given the similar lack of electronic records in other institutions.

An important element of these results is the inclusion, rather than the exclusion, of one particular kind of record: noncirculating items. There is no code in OCLC WorldCat to identify noncirculating ma-

TABLE 1	
Conspectus Subject Categories	
1.	Agriculture
2.	Anthropology
3.	Art and Architecture
4.	Biological Sciences
5.	Business and Economics
6.	Chemistry
7.	Computer Science
8.	Education
9.	Engineering and Technology
10.	Geography and Earth Sciences
11.	History and Auxiliary Sciences
12.	Invalid or unknown
13.	Language, Linguistics, and Literature
14.	Law
15.	Library Science, Generalities, and Reference
16.	Mathematics
17.	Medicine
18.	Music
19.	Performing Arts
20.	Philosophy and Religion
21.	Physical Education and Recreation
22.	Physical Sciences
23.	Political Science
24.	Psychology
25.	Sociology

terials. Identifying all possible reference materials and checking them with the local catalog is prohibitively laborious. As with most libraries, the noncirculating books represent a small proportion of the total collection. The main reference collection and the reference collections of all the branches combined represent approximately two percent of the total items studied, an amount sufficiently small to ignore without skewing the data, analyses, or results.

OCLC numbers were used to collect LCC numbers and map them to conspec-

		1998			
Book	**–1997 (before study)**	**–2002 (during study)**	**2003– (after study)**	**Total Circulations in System**	**Collection and Calculation Implications**
X	4 circs	0 circs	3 circs	7 circs	Inadvertently included (has 2+ circs and last circ'd in 2003; impossible to distinguish 2003 and pre-1998 circs from circs in study period)
Y	0 circs	12 circs	1 circ	13 circs	Correctly included (demonstrates why books with 2+ circs and last circ'd in 2003 cannot be excluded)
Z	2 circs	4 circs	0 circs	6 circs	Correctly included (counts 2 circs too many, but system cannot distinguish individual circ dates)

TABLE 2
Circulation Data Collection and Calculations

tus subject categories and to provide a picture of UCB's holdings by subject. Total monographic holdings at UCB were 970,780 titles.
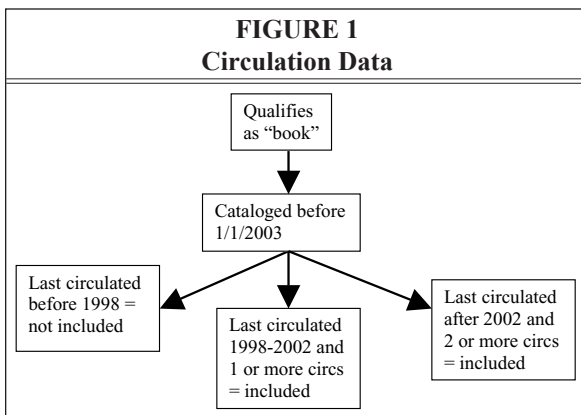
Circulation Data

The authors collected bibliographic information for every book, as defined above, that circulated at UCB from January 1998 through December 2002, the final complete year for which data were available at the beginning of this project. These dates coincide with available ILL borrowing data as described below. The authors collected the bibliographic record for every title (as opposed to every copy) that had one or more circulations in the time period defined, or two or more circulations if the most recent circulation was in 2003. (See figure 1 for a flowchart demonstrating inclusion of circulation data.) For example, an item that met the criteria for a book but last circulated before 1998 would not be included. An item that met the criteria for a book and last circulated some time between 1998 and 2002 was included. An item that met the criteria but last circulated in 2003 was only included if it had at least two circulations because at least one circulation occurred outside the window of the study. OCLC numbers were used again to collect LCCs, which were mapped to conspectus subject categories to get a picture of which subjects circulated, and in what proportions, at UCB during the period of study.

The UCB circulation system does not track all action dates on every item; it collects *only* the most recent action date and the *total* number of circulations since the item was entered into the catalog. Table 2 demonstrates how circulation data were collected and calculated.

Because the number of items resembling book X is likely to be very small, the authors elected to leave

FIGURE 1
Circulation Data

them in the data to be analyzed. Circulation statistics have been collected by the UCB catalog starting in 1995, so any books such as book Z that were published before 1995 and circulated in our time period will have three years of extra circulation tallies. The total number of circulations was 1,638,740.

Interlibrary Loan Data

The authors collected ILL borrowing data for UCB from 1 January 1998 to 31 December 2002. The starting date was controlled by implementation of the department's ILL tracking software, which provided access to the borrowing data, and the ending date was the last full year for which complete data were available at the time the project began. Data were included only for requests for books as defined above; cancelled requests were excluded from the data sample. OCLC numbers attached to the borrowing records were used to identify the LCCs for the requested titles and mapped to conspectus subject headings in order to analyze the subject dispersion of the borrowing requests. During the period of study, borrowers were not required to provide their department affiliation or academic status (faculty, staff, and so on), and very few volunteered the information. Records that included department and status accounted for less than ten percent of the total requests. This small amount does not provide sufficient data for meaningful analysis. Those fields are now mandatory at UCB to facilitate future study.

It is important to mention that UCB is a member of a statewide consortium with a shared online catalog. The consortium is composed of several library systems in the state, representing the majority of the academic libraries in Colorado and Wyoming, and many of the largest public library systems in Colorado. It is via this unified catalog that a large amount of in-state ILL borrowing requests are handled. However, lending and borrowing data are not archived in the statewide catalog, making analysis of in-state borrowing

requests impossible. Therefore, one can conclude that the ILL data collected for this particular study represent materials sufficiently rare that no other member of the consortium can provide a copy. The total number of ILL borrowing requests for the period of study was 22,064.

Results and Discussion

The data collected for this study yielded many valuable results that are too numerous to discuss in full in this paper. Based on previous studies reported in the literature, the elements selected for discussion are:

- Overall holdings
- Average transactions per item
- Percentage of items circulated in a given subject collection
- Ratio comparing ILL requests to holdings in a subject area

Overall Holdings

The monographs examined in the study were broken down into conspectus subject categories. (See table 3 for holdings figures.) The largest collections are disciplines with very high publishing output.[17] In addition, literature and history remain fields that rely very heavily on the research monograph, which may contribute to the large size of the collections.[18–20] In contrast, though UCB has a very strong chemistry program, the collection is more dominated by serials and electronic resources than monographs. UCB has an anthropology program. The small size of the collection is likely a result of the small monographic publishing output in the field of anthropology.[21] There is no sport science program at UCB. Given the enormous popularity of various kinds of sports and physical exercise in the Boulder and Colorado community, it is likely that most of this collection supports personal or popular use rather than academic discourse. In addition, a statewide reciprocal borrowing policy, whose statistics could not be tracked at the level of the data used for this project, make it more likely that such popular subjects in

| | TABLE 3 | |
| | **Number of UCB Holdings in WorldCat by Subject** | |
Rank	Conspectus Subject Category	Holdings
1	Language, Linguistics, and Literature	193,781
2	History and Auxiliary Sciences	126,797
3	Business and Economics	85,973
4	Engineering and Technology	58,377
5	Philosophy and Religion	50,024
6	Art and Architecture	46,763
7	Sociology	43,437
8	Medicine	36,501
9	Political Science	35,764
10	Invalid or unknown	35,065
11	Library Science, Generalities, and Reference	33,327
12	Education	33,314
13	Physical Sciences	28,497
14	Biological Sciences	26,234
15	Mathematics	21,145
16	Music	18,476
17	Geography and Earth Sciences	16,550
18	Psychology	13,406
19	Computer Science	12,958
20	Performing Arts	11,921
21	Law	11,721
22	Chemistry	9,504
23	Anthropology	7,688
24	Physical Education and Recreation	7,067
25	Agriculture	6,490

Average Transactions per Item

The average transactions per item represents the average number of times each book circulated. This average represents the total number of circulations within a given subject area, divided by the number of circulated monographs in that subject area. For example, if 100 books in a given subject area circulated, and those same 100 books tallied a total of 600 circulations, the average circulation per title would be 6.0. This figure reveals the subject areas that receive particularly heavy or light *use*, regardless of collection size. With a few possible exceptions, only books that circulated during the period studied were included. However, the time frame for the circulations of these books could not be defined because of limitations in the circulation system. This study examined the total number of transactions from the origin of the circulation system in December of 1994. (See table 4.)

High average transactions per item could demonstrate an extremely targeted collection or simply an extremely small collection in a popular area. Explanation for the high number of transactions per item will require further study. Though the undergraduate enrollment in computer science was not exceptionally high, it is likely that circulation was increased due to the broad applicability of computer science in many other degree programs, particularly in engineering, which has one of the highest undergraduate enrollment numbers in the university. Although the undergraduate enrollment in sociology for the period of this study was fewer than 600 students, the materials in a sociology collection have much broader applicability in all of the social sciences and in women's studies, which may have contributed to the high circulation in this area. The art and architecture programs at UCB have very large undergraduate enrollments, which may partially account for such high circulation figures.

the Colorado community would circulate via that borrowing policy. The collection in agriculture is the smallest of all subject collections at UCB. It is reasonable to expect such a small collection because there is no agriculture program at UCB.

A low average could indicate various things. It could simply be the result of not having an academic program in the area, so there is less circulation in support of research and study. It also could mean that a collection has increased in obsolescence and is in need of weeding. It is difficult to interpret the meaning of this average without further study of user behavior and collection decision making. Business and economics constitutes the third largest monographic collection, so the small transaction average is a little surprising. This is most likely the result of the speed with which business materials become obsolete and the difficulty—and sometimes inadvisability—of weeding quickly enough to keep up with the pace of obsolescence in the field. As mentioned above, the law collection at UCB is separate, which almost certainly accounts for the low average transactions in law. The

	TABLE 4			
	Average Number of Transactions per Item by Subject			
Rank	Conspectus Subject Category	Circulation Transactions	Circulating Items	Transactions per Item
1	Music	53,855	7,230	7.4
2	Computer Science	35,378	5,202	6.8
3	Sociology	106,724	17,809	6.0
4	Physical Education and Recreation	14,432	2,409	6.0
5	Art and Architecture	106,186	17,962	5.9
6	Anthropology	19,424	3,331	5.8
7	Psychology	31,281	5,376	5.8
8	Geography and Earth Sciences	31,473	5,552	5.7
9	Engineering and Technology	108,834	19,712	5.5
10	Mathematics	45,037	8,247	5.5
11	Performing Arts	25,106	4,661	5.4
12	Physical Sciences	52,423	9,786	5.4
13	Medicine	73,555	13,846	5.3
14	Agriculture	11,109	2,145	5.2
15	Biological Sciences	40,308	7,961	5.1
16	Language, Linguistics, and Literature	280,667	56,631	5.0
17	History and Auxiliary Sciences	230,262	46,515	5.0
18	Philosophy and Religion	91,324	18,696	4.9
19	Chemistry	13,011	2,775	4.7
20	Business and Economics	102,587	23,027	4.5
21	Political Science	52,108	11,745	4.4
22	Law	15,929	3,638	4.4
23	Invalid or unknown	44,707	10,419	4.3
24	Education	37,425	8,870	4.2
25	Library Science, Generalities, and Reference	15,595	4,972	3.1

	TABLE 5			
	Percentage of Collection Circulated by Subject			
Rank	**Conspectus Subject Categories**	**Holdings**	**Circulating Items**	**% Items Circulated**
1	Anthropology	7,688	3,331	43.3%
2	Sociology	43,437	17,809	41.0%
3	Computer Science	12,958	5,202	40.1%
4	Psychology	13,406	5,376	40.1%
5	Music	18,476	7,230	39.1%
6	Performing Arts	11,921	4,661	39.1%
7	Mathematics	21,145	8,247	39.0%
8	Art and Architecture	46,763	17,962	38.4%
9	Medicine	36,501	13,846	37.9%
10	Philosophy and Religion	50,024	18,696	37.4%
11	History and Auxiliary Sciences	126,797	46,515	36.7%
12	Physical Sciences	28,497	9,786	34.3%
13	Physical Education and Recreation	7,067	2,409	34.1%
14	Engineering and Technology	58,377	19,712	33.8%
15	Geography and Earth Sciences	16,550	5,552	33.5%
16	Agriculture	6,490	2,145	33.1%
17	Political Science	35,764	11,745	32.8%
18	Law	11,721	3,638	31.0%
19	Biological Sciences	26,234	7,961	30.3%
20	Invalid or unknown	35,065	10,419	29.7%
21	Language, Linguistics, and Literature	193,781	56,631	29.2%
22	Chemistry	9,504	2,775	29.2%
23	Business and Economics	85,973	23,027	26.8%
24	Education	33,314	8,870	26.6%
25	Library Science, Generalities, and Reference	33,327	4,972	14.9%

lower average transactions in education are likely the result of the separate departmental education collection. Education researchers and students use that collection often, and that use may result in fewer circulations from the main library collection. The lowest average transactions were in library science, generalities, and reference. In this subject, the low average could simply be an indicator that noncirculating reference materials are clustered in this subject. Another potential reason for such a low average could be that many sources are used in-house by librarians and not checked out even when they are used. In addition, many reference materials are clustered into this subject, and even if they are circulating titles, patrons may find it easier to copy or note down the information rather than check out the book for detailed study.

Percentage of Items Circulated

The percentage of items circulated rep-

resents the portion of the total collection in that subject that circulated in the five-year window of this study. For example, a measurement of 50 percent indicates that exactly half the monographs in the subject circulated during the study. This percentage did not reflect the popularity of any one title. Each title in this category was unique; thus, an item that circulated 100 times would be counted only once for this calculation, allowing a measurement of use that was not inflated by any particular popular title. In addition, multiple copies were counted only once, so this was an indication of circulating titles measured against held titles; any title with many owned copies would not falsely inflate the measurement of collection use. This percentage could represent a number of qualities of the collection, including the possibility that a highly circulating collection was very targeted, or very broadly applicable, or aggressively weeded. It was impossible to identify the exact cause of these calculations without further study. Overall, it is unlikely that usage of online materials had any measurable effect on these numbers, either high or low, because this study was limited to monographs and the UCB collection holds only 15,000 electronic books, a mere 1.5 percent of the total monograph collection in the library. Overall, 33 percent of the UCB collection circulated one or more times between January 1998 and December 2002. (See table 5 for information on holdings, items circulated, and the percentage of items circulated.)

It was noticeable that the anthropology collection was substantially smaller than other top circulating collections. Psychology and sociology also were extremely popular. It is likely that the statewide reciprocal borrowing policy inflates the circulation of such subjects that are popular in the greater Colorado community.

Actual circulation numbers of most English and American literature were generally high, despite the fact that this conspectus subject category overall had one of the lowest circulation percentages.

Chemistry also had a generally low circulation, which could be partially explained by the extreme dominance of serials in library expenditures toward chemistry at UCB and the low monographic publishing output in the field of chemistry.[22]

Ratio of Holdings to Interlibrary Loan

It was difficult to create a measurement of ILL activity that allows for reasonable comparison with holdings and circulation data. Actual numbers of ILL requests were helpful but needed to be considered in relation to the size of the library's owned collection because ILL activity in some subjects, such as literature and history, was bound to be inflated by very high publishing output. The authors have expressed ILL activity as a ratio to the UCB library system's overall holdings. This number compared unowned titles with owned titles, essentially comparing apples to oranges. However, the number does give an impression of how ILL use compares to the library system's holdings. The nearer the ratio approaches 1:1, the more borrowing occurs in that area when compared to the holdings. For example, if the library owned 200 titles in a particular subject and borrowed 100 other titles in that subject, the ratio would be 2:1, indicating that for every book borrowed, the library owns two books in the same subject. Or, for example, if the holdings to ILL ratio is 14.9:1, that indicates that for every book borrowed in that subject, the library owns approximately fifteen titles in the same subject. There was a broad range of ratios among the conspectus subject categories. Table 6 shows the holdings, numbers of ILL requests, and ratios for each subject.

Borrowing in the agriculture collection was quite small, which was logical given the absence of an agriculture program at UCB. However, despite the fact that the literature collection at UCB had one of the highest actual numbers of borrowing requests in the entire collection, the sheer magnitude of the literature collection made the borrowing ratio one of the low-

est. Perhaps the most important factor affecting borrowing in language, linguistics, and literature was the existence at UCB of a statewide consortial borrowing program whose ILL numbers were not stored and therefore not included in the data for this study. As a result, borrowing data considered in this study were only for items not held by any other library in the state. It was likely that this factor significantly reduced the borrowing numbers in fiction and literature. The lowest borrowing ratio in the overall collection, with 144.3:1, was

in library science, generalities, and reference. This ratio indicates that the library owns approximately 144 books for every title borrowed in the subject.

Data Summary

Though this study resulted in far more data than can be thoroughly described here, some overall statements can be exemplified by the data. There was little overlap between the subjects with large collections and the subjects with high percentages of materials circulated. Large

TABLE 6
Holdings : ILL Ratio by Subject

Rank	Conspectus Subject Category	Holdings	ILL Items	Holdings : ILL Ratio
1	Agriculture	6,490	713	9.1 : 1
2	Medicine	36,501	2,496	14.6 : 1
3	Physical Education and Recreation	7,067	387	18.3 : 1
4	Engineering and Technology	58,377	2,334	25.0 : 1
5	Invalid or unknown	35,065	1,353	25.9 : 1
6	Sociology	43,437	1,673	26.0 : 1
7	Music	18,476	654	28.3 : 1
8	Psychology	13,406	392	34.2 : 1
9	Anthropology	7,688	218	35.3 : 1
10	Biological Sciences	26,234	719	36.5 : 1
11	Law	11,721	307	38.2 : 1
12	Chemistry	9,504	246	38.6 : 1
13	Geography and Earth Sciences	16,550	377	43.9 : 1
14	Philosophy and Religion	50,024	1,086	46.1 : 1
15	Computer Science	12,958	279	46.4 : 1
16	Performing Arts	11,921	222	53.7 : 1
17	History and Auxiliary Sciences	126,797	2,343	54.1 : 1
18	Art and Architecture	46,763	728	64.2 : 1
19	Mathematics	21,145	327	64.7 : 1
20	Business and Economics	85,973	1,210	71.1 : 1
21	Physical Sciences	28,497	390	73.1 : 1
22	Education	33,314	447	74.5 : 1
23	Political Science	35,764	457	78.3 : 1
24	Language, Linguistics, and Literature	193,781	2,475	78.3 : 1
25	Library Science, Generalities, and Reference	33,327	231	144.3 : 1

collections, not surprisingly, were always in subjects with very high publishing outputs. Some subjects had high percentages of circulation, indicating an active collection. However, these subjects did not necessarily have high transactions per item, indicating that collections in which many of the books circulate are not necessarily collections in which those books circulate often.

Only a few subjects appeared in the top or bottom five subjects in most of the four analysis areas of overall holdings, average transactions per item, percentage of items circulated, and ratio of holdings to interlibrary loan. Those subjects were sociology, physical education and recreation, education, and library science, generalities, and reference. Although it was not among the largest or smallest collections, sociology had very high transactions per item, percentage of the collection circulated, and ratio of holdings to interlibrary loan. This indicates that sociology is an extremely active collection in all ways. This is possibly due to the fact that sociology books are relevant to many subjects outside sociology. Physical education and recreation was one of the smallest collections in the UCB library, probably partly due to the fact that there is no sport science program at UCB. Despite this small collection, however, physical education had one of the highest transactions per item, showing that this collection circulates frequently. Physical education also had among the highest borrowing ratios, showing that it is a very popular subject despite the lack of a related university program. Education was not among the largest or smallest subjects and appeared in the bottom of every other list. Education had few transactions per item, a very low percentage of the collection circulated, and few borrowing requests. Education is a subject that is not extremely active in the UCB libraries because a separate collection exists in the education department. Library science, generalities, and reference also did not appear among the largest or

smallest collections, but like education, appeared at the bottom of every other area of analysis, with a low average of transactions per item, low percentage of the collection circulated, and a very low borrowing ratio. This is likely the result of the fact that noncirculating materials are clustered in this area and that most users of these materials would use them in the library without checking them out.

Perhaps the overall most important factor demonstrated by this study is the importance of combining different sources of data for collection development decisions. Simply looking at collection size, circulation data, or ILL data creates a limited, and usually not entirely accurate, picture of a subject collection. Librarians must use multiple sources of data to create a truly useful picture of their collections. At UCB, these data are currently being used to assist in an off-site storage transfer project. For this project, the data were mapped to conspectus subject headings. In addition, the data were mapped to North American Title Count (NATC) subjects. The NATC divides the LC classification system into 700 categories, rather than 26. With 700 categories, the NATC is not useful for providing a broad analysis of a collection. However, since it is so minute in its subject areas, the NATC can provide an extremely detailed subject analysis of a collection. Librarians have used these combined data to single out areas of little and heavy use, defined by the combined circulation and ILL data. This allows the librarian to weed more heavily in some areas and less in others. The result is a collection that continues to meet the strongest need with the on-site collection.

Limitations of the Study

Although the results of this study have been beneficial to both the UCB libraries collection development and assessment processes, and the method can be applied to other academic and research libraries as a collection assessment model, there are limitations to the study methodology. One of the major limitations is the lack of

time that practitioners are able to dedicate to data gathering and the limited statistical expertise available for data analysis. Without the assistance of the research scientists at the OCLC Office of Research, the UCB libraries would not have been able to analyze their WorldCat holdings in order to compare the circulation statistics and ILL borrowing data by subject.

This study does not include statistics of materials that are used in the library, but not checked out, because the UCB libraries do not have the resources to collect these data. The circulation statistics and ILL borrowing data analyzed reflect only a five-year period. This is a small percentage of the many years of circulation statistics and ILL borrowing data that have been accumulated by the UCB libraries.

Since the inception of this study, the UCB librarians have decided to capture the date the catalog record was created to track the acquisition process and to analyze this information with the circulation statistics and ILL borrowing data on a quarterly basis. As discussed in the methodology, the UCB libraries did not capture annual circulation data. After the data from this study were shared with the UCB librarians, a decision was made to reconfigure the integrated library system (ILS) tracking statistics to capture the number of circulations/item/year beginning in January 2004. In January of 2005, these data were captured for use in the off-site storage transfer project. In addition, discussions are under way to configure the ILS to collect circulation statistics by subject areas identified by the bibliographers.

The identification of ILL borrowers by discipline and status was the original intent of the study. Because ILL request forms did not require discipline and status, less than ten percent of the ILL requests in the study period included this information; therefore, there was insufficient data to include in this study. As a result of this, the ILL request forms have been revised to require discipline and status.

Although it is often difficult to transition research into practice and to change current decision-making processes, the UCB librarians have implemented changes to the libraries' data collection methods based on the methodologies and results of this study.

Future Research and Conclusion

In addition to the above-mentioned changes made to the UCB libraries' data collection methods, the results of this research have provided empirical data for collection management and remote storage decision making at UCB. The results also could be utilized for the development of qualitative assessments, such as interviews with faculty, as suggested by Sonia Bodi and Katie Maier-O'Shea.[23] The circulation statistics and ILL borrowing requests could be used to calculate the obsolescence of a book based on its publication date and the increase or decrease of the number of circulations during a five-year period.

An analysis of the publishers of the items circulated and requested through ILL may provide librarians sufficient information to make collection management and remote storage decisions in a programmatic way. This could be accomplished by identifying the highly circulated and requested titles by publisher and subject. These could be compared to the titles owned by the library to determine which titles would be likely candidates for remote storage.

Identifying the languages, other than English, and subjects with high circulation and ILL borrowing requests also would support collection management and remote storage decisions. Although the criteria used for the acquisition and the methods used to identify and retrieve journal articles differ from those for books, the identification of journal titles and subject areas that are accessed through the library and requested through ILL would provide empirical data for the selection of database and online services.

The UCB librarians will gather and analyze the circulation statistics and the

ILL borrowing data on an annual basis and present the analyses and findings to the subject bibliographers. By tracking the date the catalog record was created, it will be possible to determine how soon after cataloging a title is circulated. The analysis of circulation and ILL borrowing data by subjects, as suggested by Hamaker, and borrower discipline and status will help the librarians make informed collection decisions based on users' borrowing patterns.

It is essential that collection development decisions include the consideration of data such as those gathered for this study. Usage data are even more important in the light of remote storage facilities and the attendant collection storage decisions that have been adopted by many U.S. libraries. Although the literature includes many studies utilizing circulation statistics or ILL data, very few published studies investigate the intersection of circulation, interlibrary loan, and holdings, an intersection that can provide very fruitful guidance for subject librarians.

Notes

1. Sonia Bodi and Katie Maier-O'Shea, "The Library of Babel: Making Sense of Collection Management in a Postmodern World," *Journal of Academic Librarianship* 31, no. 2 (2005): 143–50.

2. Chuck Hamaker, "Time Series Circulation Data for Collection Development or: You Can't Intuit That," *Library Acquisitions: Practice and Theory* 19, no. 2 (1995): 191–95.

3. Charles B. Osburn, "Collection Evaluation and Acquisitions Budgets: A Kaleidoscope in the Making," *Journal of Library Administration* 17, no. 2 (1992): 3–11.

4. Dennis P. Carrigan, "Data-guided Collection Development: A Promise Unfulfilled," *College & Research Libraries* 57 (1996): 429–37.

5. Richard L. Trueswell, "Some Behavioral Patterns of Library Users: The 80/20 Rule," *Wilson Library Bulletin* 43, no. 5 (1969): 458–61.

6. Barbara Hoffert, "Book Report, Part 2: What Academic Libraries Buy and How Much They Spend," *Library Journal* 123, no. 14 (1998): 144–46.

7. Chuck Hamaker, "Time Series Circulation Data for Collection Development."

8. Justin Littman and Lynn Silipigni Connaway, "A Circulation Analysis of Print Books and E-books in an Academic Research Library," *Library Resources and Technical Services* 48, no. 4 (2004): 256–62.

9. Mike Day and Don Revill, "Towards the Active Collection: The Use of Circulation Analyses in Collection Evaluation," *Journal of Librarianship and Information Science* 27, no. 3 (1995): 149–57.

10. George S. Bonn, "Evaluation of the Collection," *Library Trends* 22, no. 3 (1974): 265–304.

11. William Aguilar, "The Application of Relative Use and Interlibrary Loan in Collection Development," *Collection Management* 8 (1986): 15–24.

12. Terry R. Mills, *The University of Illinois Film Center Collection Use Study* (Urbana: Univ. of Illinois, 1981).

13. John N. Ochola, "Use of Circulation Statistics and Interlibrary Loan Data in Collection Management," *Collection Management* 27 (2002): 1–13.

14. Carrigan, "Data-guided Collection Development," 432.

15. Research Libraries Group, "A Brief History of the RLG Conspectus," Web page. Available at http://www.rlg.org/conspechist.html. [Accessed 23 July 2004].

16. Anna H. Perrault, "Library Groupings." *Global Collective Resources: A Study of Monographic Bibliographic Records in WorldCat*, Report of a study conducted under the auspices of an OCLC/ALISE 2001 Research Grant. Available at http://www.oclc.org/research/grants/reports/perrault/intro.pdf.

17. Dave Bogart, ed., *The Bowker Annual Library and Book Trade Almanac*, 48th ed. (Medford N.J.: Information Today, Inc., 2003), 495.

18. Charlene Kellsey and Jennifer E. Knievel, "Global English in the Humanities? A Longitudinal Citation Study of Foreign Language Use by Humanities Scholars," *College & Research Libraries* 65 (2004): 194–204.

19. Jennifer Wolfe Thompson, "The Death of the Scholarly Monograph in the Humanities? Citation Patterns in Literary Scholarship," *Libri* 52 (2002): 121–36.

20. John Cullars, "Characteristics of the Monographic Literature of British and American Literary Studies," *College & Research Libraries* 46 (1985): 511–22.

21. Bogart, *Bowker Annual Library and Book Trade Almanac,* 496.

22. Ibid., 496.

23. Bodi and Maier-O'Shea, "The Library of Babel."