

Report of the Collection Development Executive Committee Task Force on Print Collection Usage Cornell University Library

**Submitted to the Collection Development Executive Committee
October 22, 2010 (revised November 22, 2010)**

Task Force Charge

Reporting to Collection Development Executive Committee (CDExec), the Task Force on Print Collection Usage will conduct a wide-ranging study of the use of the circulating print collections of the Cornell University Library. The Task Force will analyze circulation data (and other measures of use, as appropriate) for recent acquisitions, as well as historical trends in collection use at Cornell, for the full range of subject areas. The group will consider usage in the context of the various disciplines and their representation at Cornell, taking factors such as language and potential readerships into account. The Task Force will work with individual selectors to understand collection use in relation to research needs in their subject areas.

The Task Force will prepare a report, to be submitted to CDExec by July 2, 2010, proposing an intellectual framework for thinking about collection “use” in the 21st-century research library, synthesizing the group’s analysis of the CUL usage data, and providing recommendations as to how usage should inform selection and allocation decisions. The Task Force will monitor ongoing strategic planning at the University and address the implications of Cornell’s emerging academic priorities for the CUL collections. The Task Force will report interim findings to CDExec at the Committee’s first meeting in April 2010.

Membership: Rich Entlich, Greg Green, Peter Hirtle, Steve Rockey, Don Schnedeker, Patrick Stevens, Kornelia Tancheva, Kizer Walker (chair). Jim Alberts served on the Task Force from November 2009 through March, 2010.

Acknowledgements

In addition to the work of the task force members, Lydia Pettis, Joanne Leary, and Peter Hoyt contributed their considerable expertise in assembling the data and making it available.

Greg Lawrence prepared the graphs on the circulation of Agriculture monographs and on monograph usage in the Department of Horticulture and the graduate field of Horticultural Biology (figures 31-34).

Executive Summary

This report begins with a discussion of the rationale for examining the use of CUL's print monograph collections, situating the study in the context of both the current budget-related changes at Cornell and broader shifts in the nature of research library collections (e.g., the shift to electronic, demand-driven acquisitions, collaborative collection development, etc.). The task force then presents several assumptions about print collections and academic research that have guided the study and enumerates several factors that have complicated the analysis of the circulation data.

Data about CUL acquisitions since 1990 is presented, showing the distribution of print monographs in the collection by subject, by language of publication, and by unit library. The task force then considers data from three types of usage reports: a cumulative tally of circulation, by publication year, of monographs in the CUL collection published 1990 to 2010; a count of the annual circulation of books acquired by CUL in 2001; and a "snapshot" of circulation on one day in April 2010. The report includes representative examples from the various reports, often using the Olin Library collection before it was re-mapped in July 2010 to illustrate points about the data; complete data files are available online.

Highlights from the data analysis include:

- Approximately 45% of print monographs in the CUL collection published since 1990 have circulated at least once to date; approximately 55% of these books have never circulated.
- Books in 337 languages were acquired from 1990 to 2010 throughout CUL. 55% of these were in English. 61% of Olin Library's English-language books acquired since 1990 have circulated.
- Circulation of monographs published since 1990 has tended to increase gradually for 12 years, at which point the use of new volumes tends to level off.
- For books published in 2001, the average unique circulation CUL-wide holdings was 20% by the end of 2002, 30% by the end of 2005, and 35.5% by end of 2009 (the unit libraries differ significantly with regard to both the rate at which previously non-circulating titles continue to be checked out over time, and the point at which use of new volumes starts to level off).
- Most books in circulation on April 19, 2010, were charged to graduate students, who accounted for 34% of the total charges. Faculty had out another 23.6%. Undergraduates had out only 10.7% of the books charged – 16,744 books in total or an average of about one book per undergraduate student in the Cornell population (compared to approximately 8 books, on average, for graduate students and about 13 per faculty member).
- Examination of books charged to users affiliated with a sample of academic programs reveals use of books in a wide range of LC classifications. There are also significant differences among academic programs in the degree to which they are using books from beyond their fields.

The library in the research university has traditionally aspired to build a collection that would satisfy any potential research need; that some portion of the collection would remain indefinitely latent has generally been accepted as the condition for meeting the needs of scholarship. What significance the Library and the University should assign to non-circulating material in today's academic context is far from clear, however. If half of CUL's monograph purchases of the past twenty years have circulated, is that a lot or a little? Precious resources are being spent to purchase, house, and preserve these books, but to what extent should this be regarded as misspent funds and to what extent as investment in a strategic reserve? The answer will surely vary by field and by the intended readership for particular segments of the collection. Factors such as language of publication can place distinct limits on the pool of potential users and any meaningful measure of usage must take the size of the user population into account.

The report provides the following recommendations to CUL administration:

1. Integrate monograph usage data and its assessment into CUL's collection development operations at the administrative level.
2. Make past monograph usage, understood in context, one facet in decision making about future CUL acquisitions and investment in the collection.
3. Make monograph usage data accessible to subject selectors.
4. Let a more detailed understanding of usage inform where and how we house books.
5. Consider how usage data for the academic programs relates to the size of these units and the Library investment in the collections that support these fields.

In light of the many factors affecting collection use, the task force also recommends against drawing certain conclusions from the data:

- High or low circulation rates should not be attributed to a single straightforward cause, particularly in light of wide variation in the role of print monographs in different disciplines.
- The Library should not adopt specific across-the-board targets for the circulation rate of print monographs acquired for the collection.
- The Library should not halt or diminish acquisitions in particular non-English languages absent a detailed understanding of language distribution among the disciplines and across the broad patron base on campus.

Context and rationale

For reasons laid out in the “complicating factors” section below, this report restricts its view of print monograph usage to charges (excluding renewals) of books in the circulating collection.¹ This document considers data from three types of usage report: a cumulative tally of circulation, by publication year, of monographs in the CUL collection published 1990 to 2010; a count of the annual circulation of books acquired by CUL in 2001; and a “snapshot” of circulation on one day in April 2010. The narrative report that follows includes representative examples from the various reports, often using the Olin Library collection to illustrate points about the data. The complete data files are available online in a series of spreadsheets²; the data description document in Appendix 1 provides a breakdown and explanation of the various data reports.

Several factors gave rise to the present study. Most immediately and most locally, Cornell's budget situation in the wake of the global economic crisis led to the first-ever cut to the Library's endowed materials budget in fiscal year 2009-10. Further, recurring reductions to the endowment payout that supports the collection continue to curtail collection development. At a time when it is ever more critical that we prioritize our acquisitions, data about how our collections are used can help inform collection

¹ Please note that unless specifically mentioned otherwise, “use” and “usage” in the present report refer to check-out circulation of print monographs.

² The files are linked from the Collection Use Study Task Force page on CUL's Confluence wiki: <https://confluence.cornell.edu/display/culcdsc/CUL+collection+use+study+%282009-10%29>.

development decision-making, both at the level of individual selections and at the level of the allocation of funds. Usage data is regularly consulted for the electronic resources in which the library invests but, up to now, such data has not been systematically examined in the case of print materials. Outside of anecdote and conventional wisdom, little has been known to date about the use of monographs in Cornell's holdings.

Along with the immediate pressures on the Library budget, the current recession has precipitated a broad "reimagining" process across the University, and this has included discussion of implementing a more focused definition of Cornell's academic program emphases. It is as yet unclear what this could mean, concretely, for teaching and scholarship on campus, but it seems essential in this context to develop a more nuanced understanding of how students and faculty in the various academic programs are using our collections and which parts of the collection serve which fields on campus.

A complex understanding of the use of the print collections also improves the prospects for wise decision-making as the Library begins to embark on plans to coordinate collection building in certain areas with our partners at Columbia University and in the context of the Borrow Direct consortium.

Well beyond the Cornell context, multiple converging factors have influenced research libraries, formerly defined by the aspiration to build comprehensive collections on the "just-in-case" model, to move, incrementally and to varying degrees at different institutions, toward a "just-in-time" model defined by immediate user demand. The emergence of e-books, which promise never to go "out of print," and new purchasing models such as patron-driven acquisition³ have accelerated this trend. In an effort to re-focus collections (and expenditures) on immediate user needs, some libraries have linked usage assessment very directly to collection development efforts. In a rather radical example, the University of Virginia Library implemented the so-called "Balanced Scorecard" method in the early 2000s, setting targets according to which monographs that circulate within the first two years after acquisition should account for at least 60 percent of the budget for monographs.⁴ Our present study is, in part, an attempt to open a dialogue about how collection usage might appropriately inform collection building in the Cornell context. The task force feels strongly that there is no simple or one-size-fits-all answer to this question.

Guiding assumptions

A few basic assumptions about the use of research library collections, and print monograph use in particular, underlie this report. Not all of these factors are quantifiable or verifiable based on circulation data alone, but the task force has found that the data does not contradict, and tends to support, these assumptions:

1. **Disciplinary context.** Use of library collections is intricately linked to specific disciplinary cultures and to the fluid status of the academic disciplines. Cultural factors affect format preferences: although reading and research practices are shifting in all areas of scholarship, it is established that use of electronic resources is still significantly more prevalent in the sciences and quantitative social sciences than in

³ On the relationship among e-books, just-in-time collection building, and the economic crisis, see Greg Raschke and Joseph Esposito, "Building collections in a bad economy," exchange on LibLicense-L, June 15-16, 2009. Archived at <http://www.library.yale.edu/~llicense/ListArchives/0906/msg00061.html>.

⁴ Jim Self, "Using Data to Make Choices: The Balanced Scorecard at the University of Virginia Library." ARL Bimonthly Report 230/231 (October/December 2003) 28-9. <http://www.arl.org/bm~doc/balscorecard.pdf>. It should be noted that UVA has quietly dropped this method as it pertains to collection development.

humanities fields.⁵ At the same time, this rather intuitive point should not be allowed to obscure the pronounced differences within these large subject groupings as regards the use of print and electronic materials. And apart from the issue of publication medium, the disciplines are still divided around the relative importance they assign to journal and monograph publications; while the reliance on journal literature is arguably greater in the sciences than in the humanities, here again it is important to keep subtle differences among the individual disciplines in mind.

Use of library collections over time is also dependent on the dynamics of ebb and flow affecting disciplines nationally and globally, as well as at particular institutions. Subjects, approaches, and even whole fields may gain or lose momentum and status, impacting demand on library collections in complex ways.

2. Publishing context. In addition to research practices bound to disciplinary cultures, the use of library print collections is determined in part by the publishing context of the respective disciplines, i.e. the proportion of electronic, hybrid, and print-only material made available by the publishers that supply the various academic fields. This factor is closely related (but not identical) to the actual proportion of library holdings in various formats, which obviously affects the potential usage of print monographs in the collection.

3. Library context. Use of library collections in general, and print monographs in particular, is influenced by the relative collection strengths at a particular institution. For example, if a library has built a core print collection in a particular area that is often used for teaching purposes, its circulation rate (the percentage of the materials that circulate, as well as the frequency of circulation) is likely to be higher than that of an extensive research collection in another area. The levels at which a library collects are closely tied to the (current and historical) state of specific disciplines and the department(s) that engage in them at the institution. It should be noted that this is, to some extent, a two-way relationship: libraries make decisions about investing in particular subjects based on the perceived status of a field on campus, but the status of the field also derives, in part, from the strength of the library collections in that area.

4. Expert user context. Regardless of format, the use of particular subsets of a library collection depends on the number of potential users in the population served by the library. Highly specialized materials are likely to be used by a small number of specialists, and thus to circulate at a relatively low rate.

5. Linguistic context. The language of publication impacts the use of library materials differently in different subject areas. Usage depends on the number of people with reading knowledge of the respective language; the size of these potential readerships is difficult to determine, varies unpredictably over time, and does not map neatly onto academic programs. There is a linkage between the expert user context and the linguistic context.

These various contextual considerations, along with the factors described below that complicate the data, suggest that conclusions drawn from the data presented here should be considered tentative. The task force feels strongly that certain conclusions should *not* be drawn, however, and these include:

- attributing a high or low circulation rate to a single straightforward cause
- setting a specific, across-the-board percentage of “acceptable” use of acquired print monographs

⁵ E.g., Roger C. Schonfeld and Ross Housewright, “The Format Transition for Scholarly Works,” *Faculty Survey 2009: Key Strategic Insights for Libraries, Publishers, and Societies* (New York: Ithaka, 2010).

- halting or diminishing acquisitions in particular non-English languages absent a detailed understanding of language distribution among the disciplines and across the broad patron base on campus.

Complicating factors

Our ability to use quantitative data in collection assessment is sharply limited by the nature of the data available to us. For many of the questions we might like to ask, the necessary data is simply not there. Where data does exist, we face a significant challenge in ensuring that factors that are being compared are truly comparable. The division of CUL into some twenty unit libraries, each with a degree of autonomy in establishing policies and procedures, almost guarantees some variability in the data. Where we were able to filter out data that could exacerbate these variations, we have done so. We are satisfied that any lingering differences in the measures analyzed here are relatively minor.

The question of collection usage is at the heart of our study, but we discovered tremendous variation in the usage data that has been collected over time. In order to make valid comparisons across units and over time, we found it necessary to restrict our analysis to monograph *circulation*. There are both definitional and system-related issues. The primary definitional problem is that what constitutes use of a print monograph is not consistent across CUL. Most unit libraries, for example, track in-library use (called “historical browses” in the Voyager system), but the years in which the units started recording this data vary. The policies for recording in-library browses are inconsistent across units and, in any case, browses are recorded by a simple counter: the browse dates are not retained. For these reasons, we have elected to concentrate on lending rather than in-library use, even though we realize that this may skew the data, especially for collections and call number ranges that include large numbers of non-circulating reference materials. We have excluded all designated non-circulating items from the circulation analyses in this report, although we recognize that some unit libraries do, in fact, lend these “non-circulating” items under certain circumstances. We recognize that circulation is an imperfect surrogate for *use* of items in the collection.

A second definitional problem relates to what constitutes a circulation transaction. Because the length of time that an item can be borrowed varies according to the nature of the material and the status of the borrower, one item might circulate many times, but actually be out of the library less than another item that has circulated less frequently. A reserve item circulates for a few hours; movies can circulate for a few days; undergraduates, graduate students, and staff all have different borrowing periods. For part of the two decades of available data, faculty had an indefinite loan term – books charged to faculty never had to be returned. These circulation periods imply qualitatively different types of use, but for the purposes of this report, we opted to count individual circulation transactions equally, regardless of the status of the patron or the length of the loan. We assume that an initial charge represents a more significant event than a renewal and, in light of the variation in loan period among patron groups, initial charges are comparable in a way that renewals are not. We have thus excluded renewals from the circulation data. Reserve transactions are included, however (there is no way to filter them out).

The third definitional problem concerns the nature of the material whose circulation we are counting. The goal of our study was to measure the use of monographs, rather than serials, and the former only in printed form. MARC coding for monographs can include electronic resources, music scores, sound recordings, and other types of material. Our study excludes all non-language material except for music

scores, which we have purposely included.⁶ While a few anomalous items might still be reflected in the sampled data, we are confident this is not enough to have a significant impact on the analysis.

Selection criteria for all reports:

- 06 - Type of record
 - a - Language material
 - c - Notated music

 - 07 - Bibliographic level
 - m - Monograph/Item
-

Increasing transfers of books from the unit libraries to the Library Annex also complicates our data. Although most Annex transfers have been of materials with historically low circulation, some items are housed in the Annex immediately upon acquisition. Some units have moved large numbers of items to the Annex, others relatively few. It is difficult to gauge the effect of housing a volume in the Annex on its tendency to circulate, and data on Annex material may not be entirely comparable with data from the units. Separating Annex materials out prior to analysis is also problematic: since these are largely items with historically low circulation, excluding them artificially boosts the apparent usage of the unit's remaining onsite collection. We elected to include Olin materials housed at the Annex in the circulation data for the 1990-2010 Olin materials because, for the early years in that range, significant numbers of low-circulating items have been moved to the Annex and to exclude them would skew the data. In contrast, the data for monographs purchased in 2001 (fewer than 10% of which are housed at the Annex) excludes all Annex materials. The "snapshot" data, which tracks circulation of materials regardless of publication or acquisition date, includes items housed in the Annex.

As for the system-related issues, available circulation data differs according to whether it was captured during the era of the NOTIS library management system (1990-2000) or during the Voyager era (2000-present). Prior to 1990, there was no automated circulation system in use and we have no collective circulation data. Machine-readable historical circulation data begins with CUL's adoption of the NOTIS system, but this data has a number of limitations:

- There is only data for books with barcodes. Since much of the older Olin collection was not barcoded at the time, complete circulation data only exists for titles acquired after 1990 (since new titles were routinely barcoded). Items published prior to 1990 may or may not have been barcoded when NOTIS was brought online.
- The historical circulation data exists in the form of a counter. We know whether a title circulated, and how many times it circulated, but not when that circulation occurred. A 1990 book that circulated 10 times, for example, might have circulated 10 times in the first year after publication, or it might have circulated once a year for the entire decade. There is no way to distinguish.
- Practices in some unit libraries may also skew the data. For example, we discovered that the Math Library routinely charges out new titles to its "New Books" shelf. This means that almost all new books in the Math Library have at least one circulation. No other unit library appears to follow

⁶ Scores are tallied in the spreadsheets in the online appendices to this report; the narrative report itself does not address usage of musical scores.

this practice today, but there might have been similar idiosyncratic practices during the NOTIS era that also skew the data but which have been forgotten.

- Our research interests include linking circulation with demographic data. There is no such linkage for items in the NOTIS era, however. We know that a book was checked out, but we know nothing about the borrower. Thus, routine circulation to Cornell community members is counted the same as interlibrary loan circulation to external borrowers, circulation to local non-Cornell borrowers, and circulation for internal CUL purposes, such as conservation and digitization.
- When NOTIS records were moved into Voyager, all items were assigned a “create date” of May 31, 2000. Therefore, we have no hard data as to when NOTIS-era acquisitions became available to circulate. We can only assume that most newer publications were acquired approximately when they were published, though we know this to be inaccurate in some cases.

From the advent of the Voyager era, the amount of information available for analysis increases greatly. At the same time, since we have only ten years of Voyager data, and since many of the titles in Voyager were acquired relatively recently, the ability to identify long-term trends is limited. Voyager circulation data is still recorded in the counter for “historical charges,” but there is also a circulation transaction log that can detail when an individual title was checked out. It also is linked to some rudimentary demographic data about the borrower. We can know, for example, whether an item was checked out to an undergraduate, graduate student, faculty member, staff member, or one of several other special categories of borrowers. This includes internal use such as charges to the “New Books” shelf. This allows us to filter out certain types of use where appropriate. We do not know, however, in what school, department, or field of study the borrower was located.

There are anomalies in the Voyager data as well. For example, the current project to convert charged titles from “indefinite loan” to a generic faculty loan will require checking in and then checking out all books out on indefinite loan – thus skewing the circulation data for those titles in ways that will impact future studies of this type.

While the historical data from the NOTIS and Voyager systems can provide some general indication of the circulation of the library’s collection, the Task Force was interested in a much more granular analysis of collection use. Longitudinal data to support such analysis does not exist. For that reason, the Task Force implemented a “snap shot” method that provides considerable detail on the status and departmental or field affiliation, among other data points, of patrons who have books checked out at a specific moment in time.

The narrative that follows presents representative data from three types of usage report: a cumulative tally of circulation, by publication year, of CUL books published 1990 to 2010; a count of the annual circulation of books acquired by CUL in 2001; and a snapshot of circulation on one day in April 2010.

CUL monograph holdings, 1990-2010 publications

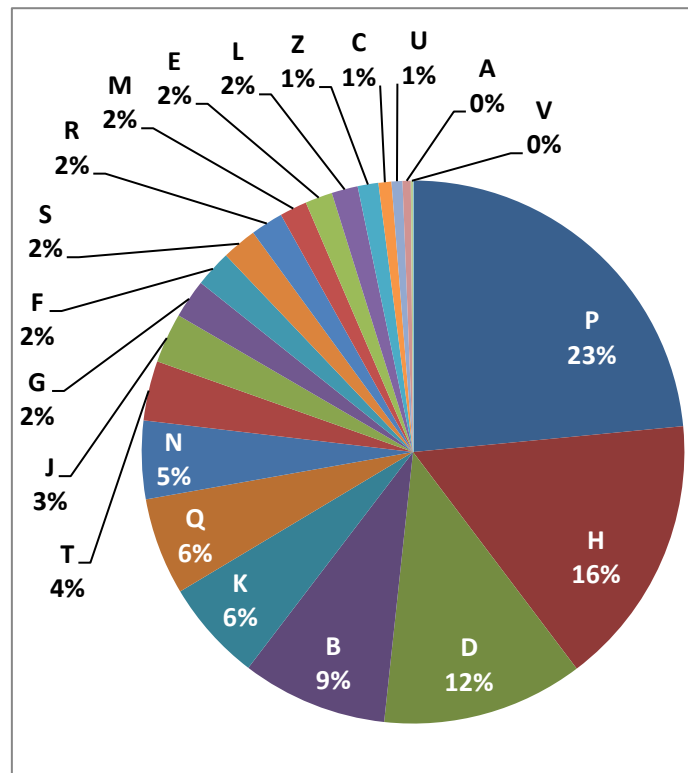
With the exception of the snapshot data, which measures circulation of the CUL print monograph collections as a whole, the circulation data presented in this report is for books acquired since 1990. Before we turn to the circulation analysis, it is essential to have a high-level picture of the collection built over this period. This section shows the distribution of 1990-2010 monographs in the collection by broad subject, by major language, and by unit library.

According to a count of holdings records added, as of mid-April 2010 CUL had acquired a total of 1,654,034 print monographs published in 1990 or later (the number is for titles, including multi-volume

works). The distribution varied considerably among subject areas as measured by the Library of Congress Classification System. The following table provides a roster of system-wide purchases by “top class” (i.e., the first letter of the LC class), in descending order of number. The graph dramatizes the preponderance of classes P, H, D (together the majority) and B, all in the realms of humanities (prominently languages and literatures), social sciences or history and an order of magnitude above the (nonetheless) strong showings in K through F.

Figure 1. Print monographs published 1990-2010 in CUL holdings, by top LC class

LC top class		vols.
P	Language and Literature	388,379
H	Social Sciences	268,291
D	World History	198,440
B	Philosophy. Psychology. Religion	143,547
K	Law	99,847
Q	Science and Mathematics	95,823
N	Fine Arts	77,022
T	Technology	58,858
J	Political Science	49,178
G	Geography. Anthropology. Recreation	38,322
F	History of the Americas	36,144
S	Agriculture	34,357
R	Medicine	32,059
M	Music and Books on Music	26,914
E	History of the Americas	26,866
L	Education	25,837
Z	Bibliography. Library Science.	20,391
C	Auxiliary Sciences of History	12,769
U	Military Science	10,755
A	General Works	8,071
V	Naval Science	2,159
[W]		3
[I]		1
[O]		1
Total		1,654,034

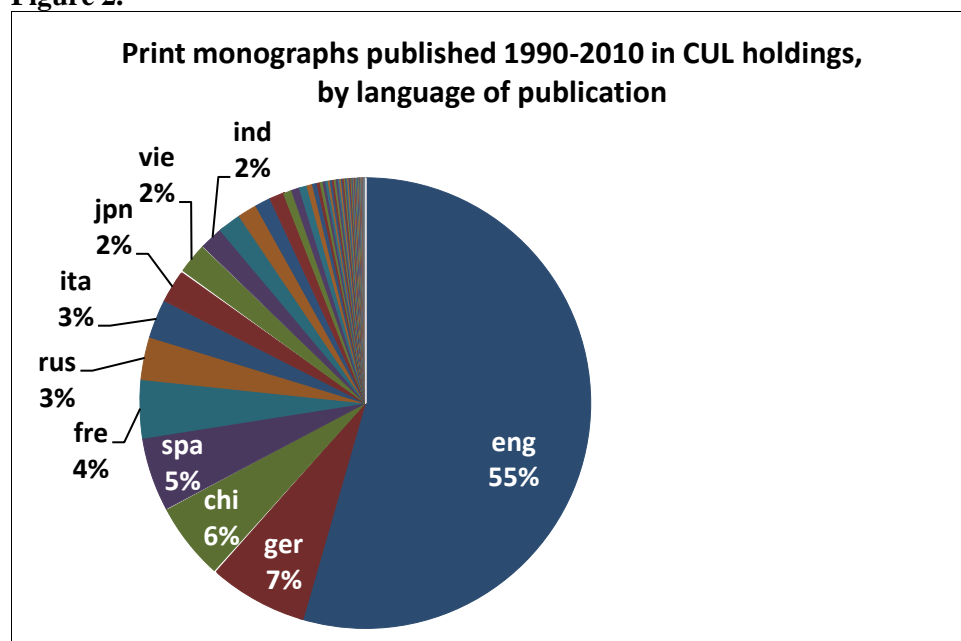


The first ten of twenty-one classes (through G) account system-wide for well above 75% of the monograph acquisitions under consideration. (The five items classed W, I and O are anomalies: there are no such classifications in the LC system.)

Language

The 1990-2010 print monographs under consideration here included books published in 337 languages. English-language books constitute 55% of the total and the first five languages taken together comprise above 75% of system-wide monograph purchases in this period. With the exception of Chinese, third in terms of numbers, the languages were major European idioms. The first ten languages account for approximately 90% of the purchases. Rounding out the second tier of five are notable purchase levels of monographs in Japanese, Vietnamese and Indonesian.

Figure 2.



Buried deeper in the statistics, books in the next ten languages (ranked 11 through 20) were nevertheless acquired in significant numbers:

Thai	25,665 volumes
Arabic	21,188 volumes
Portuguese	17,318 volumes
Hebrew	17,163 volumes
Polish	9,957 volumes
Hindi	9,323 volumes
Korean	9,239 volumes
Malay	6,843 volumes
Burmese	5,338 volumes
Dutch	3,626 volumes

Each of these languages falls short of 2% of total acquisitions of 1990-2010 monographs, but materials in all of these languages would be seen as essential to particular fields of study on campus. The rankings suggest an attribution of importance at Cornell to a diversity of Asian and Semitic as well as smaller European languages. Books in Icelandic, the chief language for CUL's Fiske Icelandic Collection, which is considered a collection of record in Anglo-American libraries, amount to just above 0.1% of the 1990-2010 monograph holdings at 2,142 volumes (for a rank of 28).

A better understanding of the world’s total available publishing output by language would add needed context to this data.

Figure 4.

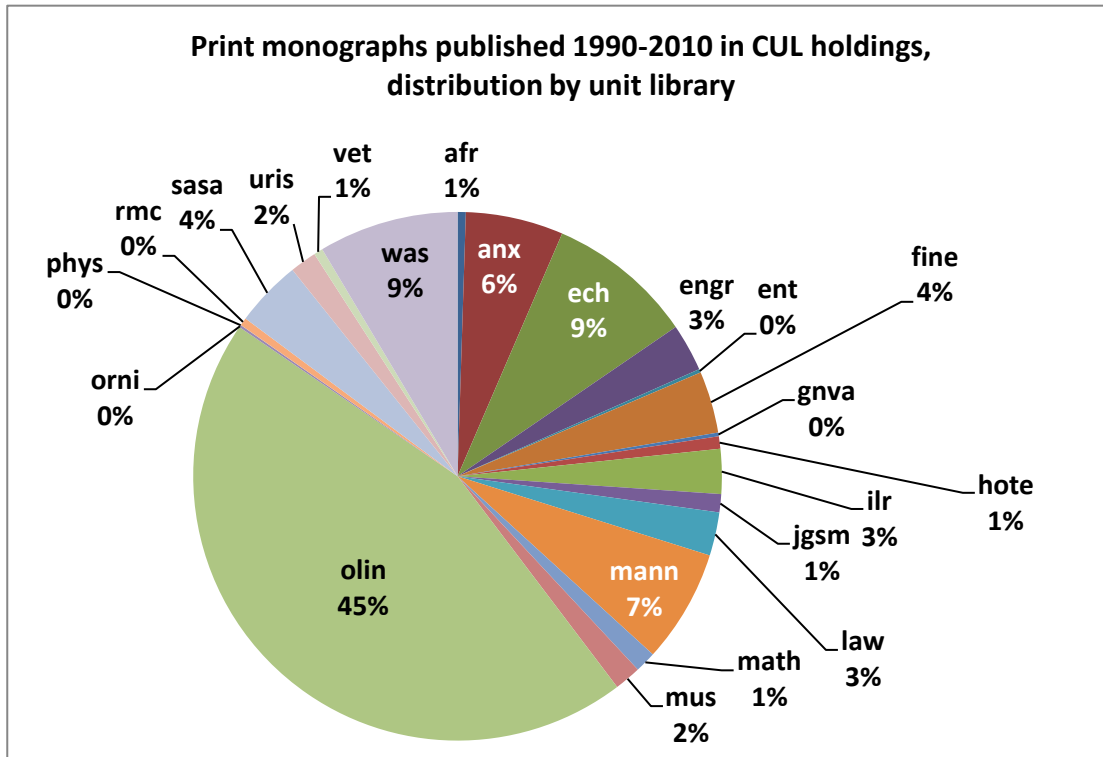


Figure 4 shows the distribution of the 1990-2010 monographs among the various CUL unit libraries. Olin Library acquired by far the largest proportion of these books – 45% – with the three Asia collections housed in Kroch Library representing the next largest segment at 22% taken together. It is notable that the Library Annex houses 6% of these relatively recent publications, nearly as large a share of the total as Mann Library. It is important to bear in mind that these percentages refer only to print monographs in the collections of the various units; the unit libraries that house large collections in the sciences and certain areas in the social sciences have a larger proportion of journals than monographs.

CUL monograph usage patterns over time

Our task force was interested in tracking trends in the circulation of the print monograph collection over time. In particular, we wanted to determine the extent to which monograph usage changes with the length of time books are held in the collection and how such changes relate to factors such as field of research. Among the questions we posed were these: How long does it take for books in various subjects to circulate? Is there a point at which books stop being “discovered” in the collection in large numbers? How does this vary by subject? In the various units, which subjects stand out as high-use and low-use among these recent acquisitions? How does language correlate with circulation for the various subjects?

Our study looked for such trends in two sets of data: circulation records for monographs in the CUL collection published between 1990 and 2010, and circulation records for monographs in the collection that were acquired in 2001. Each data set provides information that the other cannot.

1. Circulation of monographs published 1990-2010

Analysis of the 1990-2010 monographs provides a needed historical dimension to our study, and the results are revealing. However, the relative lack of detail in the NOTIS-era circulation data limits the distinctions we can draw. Because the date an item circulated was not retained before 2000, it is impossible to track a fixed set of items acquired before this date over time. We can look at everything published in 1990 and determine what percentage has circulated as of some recent date, i.e., after twenty years. We can establish the circulation percentages to date of items published in 1995 and use that as an indicator of circulation after fifteen years. If we wish to compare fifteen-year circulation to twenty-year circulation on this basis, however, it is essential to bear in mind that the circulation numbers pertain to entirely different sets of materials. Such analysis compares the circulation of the 1990 publications after 20 years with the circulation of the 1995 publications after 15 years; it does not track the use of the 1990 material, or the 1995 material, at two different points in time. Also, while most books are acquired by CUL within a few months of publication, this is not always the case. Some books published in 1993 will not have been acquired by the library until 1995 or 1996, for example, giving them that much less time in which to circulate than other books published in the same year. For the books published and acquired during the Library's NOTIS era, there is no way to know for certain which year they became available for Cornell patrons.

With those caveats in mind, one can draw some preliminary conclusions about monograph usage over time from the 1990-2010 circulation data. Figures 6 and 7 below illustrate the percentage of print monographs published in this period that had circulated at least once as of mid-April, 2010, by year of publication; figure 6 shows the trends for CUL as a whole and figure 7 graphs circulation for Olin Library.⁷

The graphs seem to show that, for the books under consideration, circulation has risen fairly steadily year by year, the longer books were in the collection, until beginning to plateau at around 12 years. For Olin books, circulation has started at a lower rate and risen more sharply than for the CUL monograph collection as a whole, but in each case, around 50% had circulated after 12 years, and books published 16 years ago, in 1994, reached the high-water mark for Olin (50.6%) and for the entire CUL collection (55.1%). The pattern we see in the graphs likely indicates that books can require an interval of several years before they are discovered in the collection, but that books that do not circulate in this interval of active discovery are significantly less likely to do so in subsequent years. There are other possible interpretations of the data, however. Since we do not know the date of circulation for books that were charged out before 2000, it is possible that most of the circulation for the older books was in the first year or two that they were in the collection; if that were the case, the lower numbers for the recent books might indicate less use of the print collection in general in recent years, rather than a leveling-out of interest in the older materials after a process of gradual discovery.

⁷ Figures 6 and 7 include Annex holdings for the publication years in question. For the Olin data, when Olin transfers to the Annex are excluded, there is an unusual bump in Olin circulation for items published 1990-1992 (between 74% and 77%), but this is due to the removal of large number of low circulation items to the Annex for those publication years.

Figure 6.

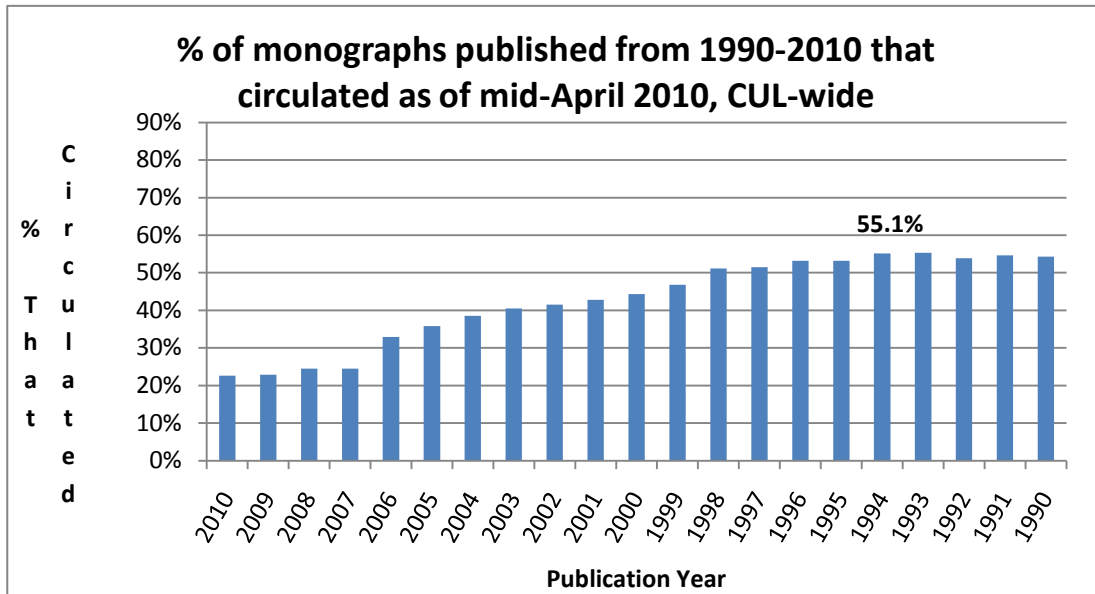
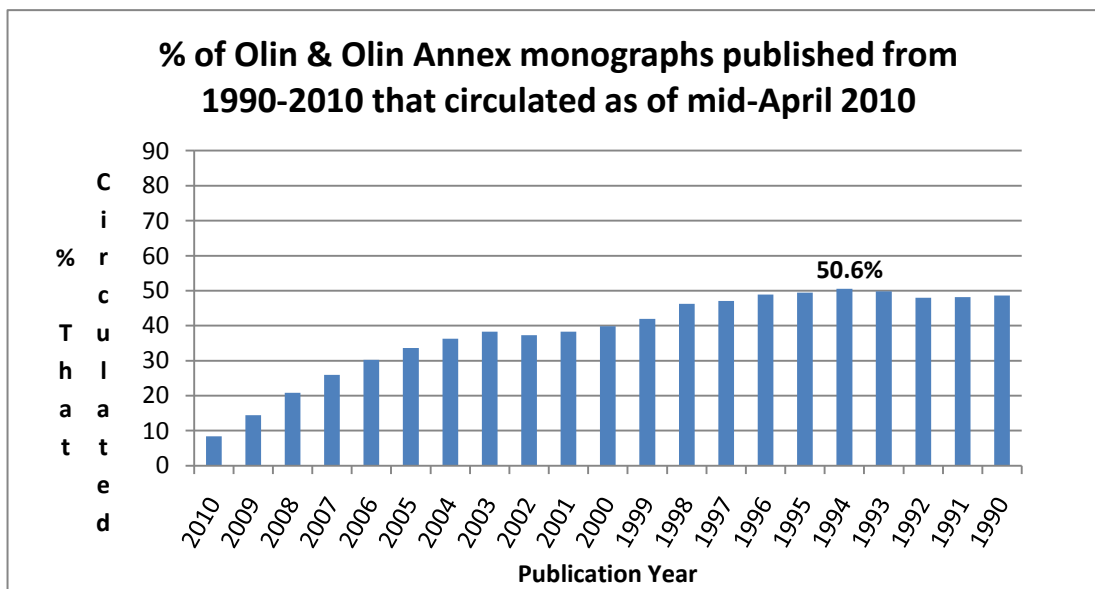


Figure 7.



Language

The cumulative circulation data for 1990-2010 monographs confirms a wide disparity between the rate of use of English-language books and books published in all other languages. Focusing on Olin Library holdings,⁸ figures 8 and 9 show the percentage, by language, of books published in this period that had

⁸ Figures 8 and 9 include Olin-Annex holdings.

circulated at least once by mid-April 2010. Figure 8 illustrates circulation of books in the twenty most heavily-represented languages in the Olin holdings. Most Olin books, of course, are in English (a volume count of 401,623 published 1990-2010), and 61% of these books have circulated. German-language books make up the next largest segment of the Olin collection; 20% of Olin's 85,454 German-language books published between 1990 and 2010 circulated. 28% of Olin's 1990-2010 French-language books, which rank fourth in holdings at 54,824 volumes, circulated in the period under review.

Figure 9 shows the most intensively-used languages in terms of the percentage of 1990-2010 Olin monographs in the respective languages that circulated at least once by mid-April 2010. Several languages of publication that are not strongly represented in the Olin collections were used relatively intensively in this sense. For instance, 34% of Olin's Turkish-language books published in these years circulated, the highest percentage of any language other than English, but this amounts to circulation of 214 out of a total of 626 books in Turkish. The use of books published in Catalan (cat) and Ancient Greek (grc) was also relatively high in terms of the percentage of circulation (542 books published 1990-2010 are coded as being in Catalan in the MARC records; 429 are coded as being in Ancient Greek). Heavy circulation of small numbers of available books in a certain language probably indicates highly targeted selection in that language; we cannot know from this data whether the apparent demand would carry over to more extensive collecting in these languages.

2. Annual circulation of monographs acquired in 2001

Because of the shortcomings of the pre-2000 data, we have only a limited ability to discern circulation trends for the 1990-2010 publications. When we restrict our analysis to Voyager-era data, several additional facets can be taken into consideration, for a much more nuanced view of monograph use, though obviously one with a shorter historical span. Beginning in 2000, we know, in most cases, precisely when an item became available to circulate, since for books received after implementation of the Voyager system, the record "create date" is normally equivalent to the date of acquisition. And unlike with the historical circulation counts, each charge recorded in the Voyager circulation transaction archive includes a time/date stamp. These two factors allow us to reliably track circulation trends over time. Moreover, we can follow our patrons' incremental discovery of particular titles over a period of years: not only can we see that a certain percentage of a given set of volumes circulated in a particular year, we can also detect when specific titles that had never circulated do so for the first time. However, if the 12-year discovery hypothesis suggested by the historical data is correct, that full cycle is not visible in this set.

The Voyager archive also permits us to track broad user community trends, since each record also includes a patron status value. Among other things, this allows us to filter out use by non-Cornell patrons, as well as internal Library circulations for special projects, including charges to unit library new books shelves, which would otherwise skew the data.

2001 is the first full year for which we have a complete set of Voyager circulation transaction archive records, and we have focused this part of our analysis on print monographs acquired in 2001. This provides over eight years of transaction data to follow.

Figure 8.

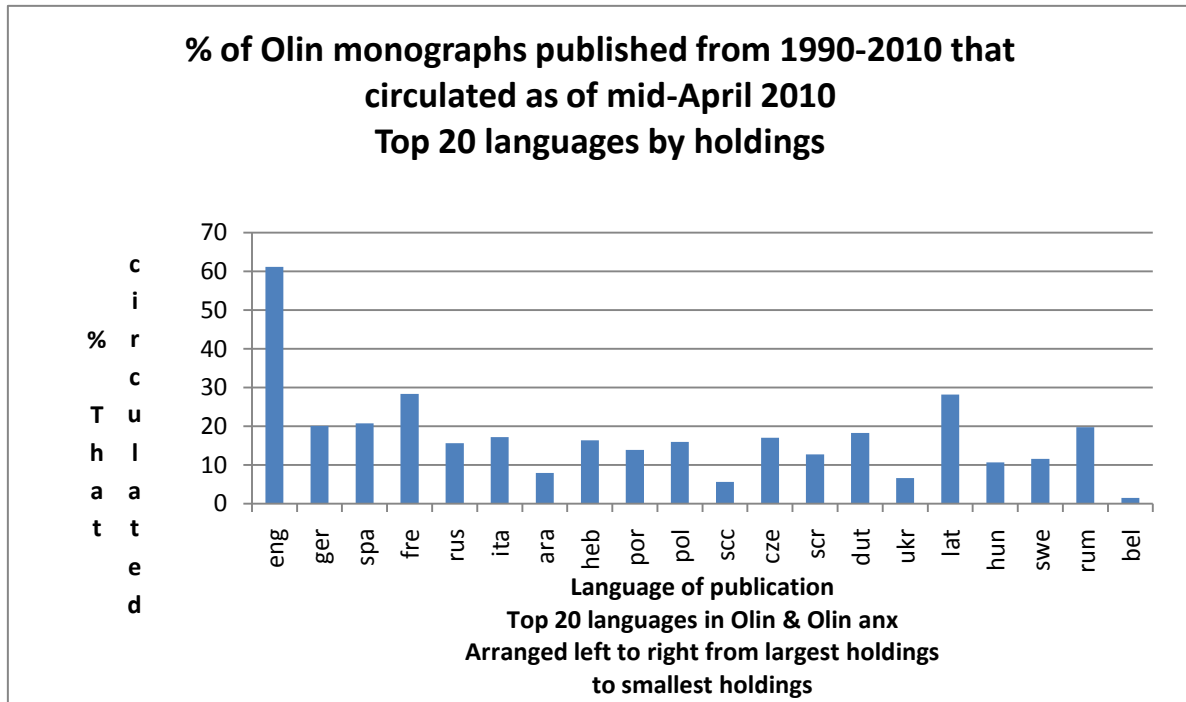


Figure 9.

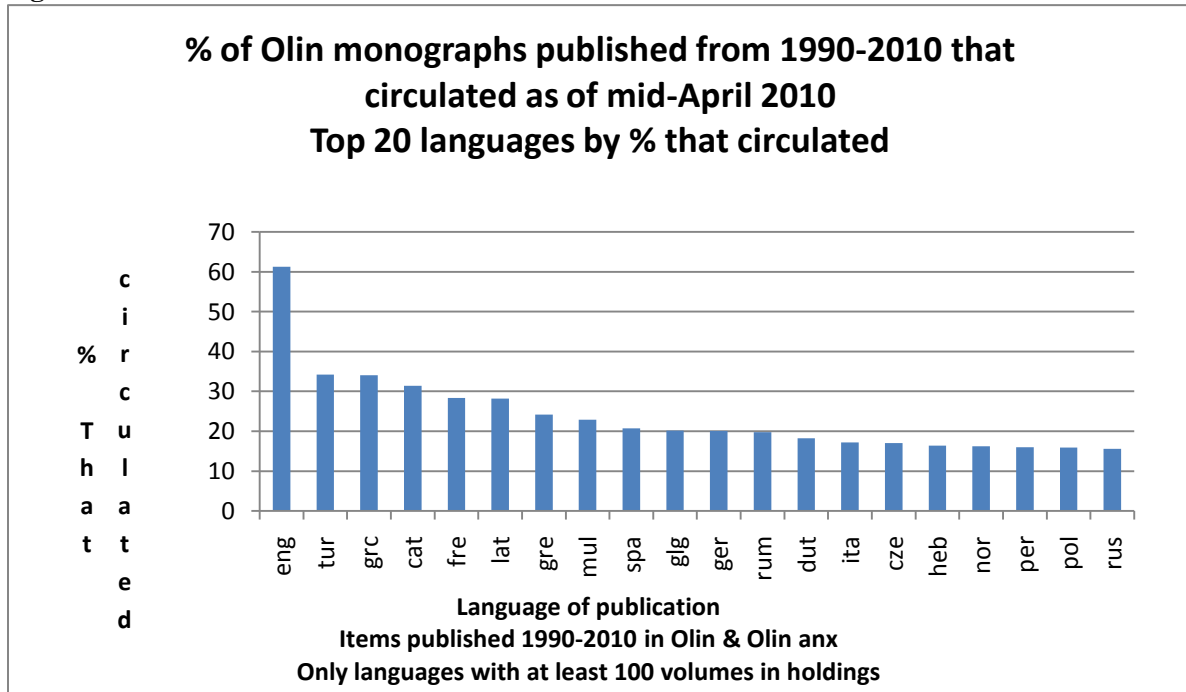


Figure 10.

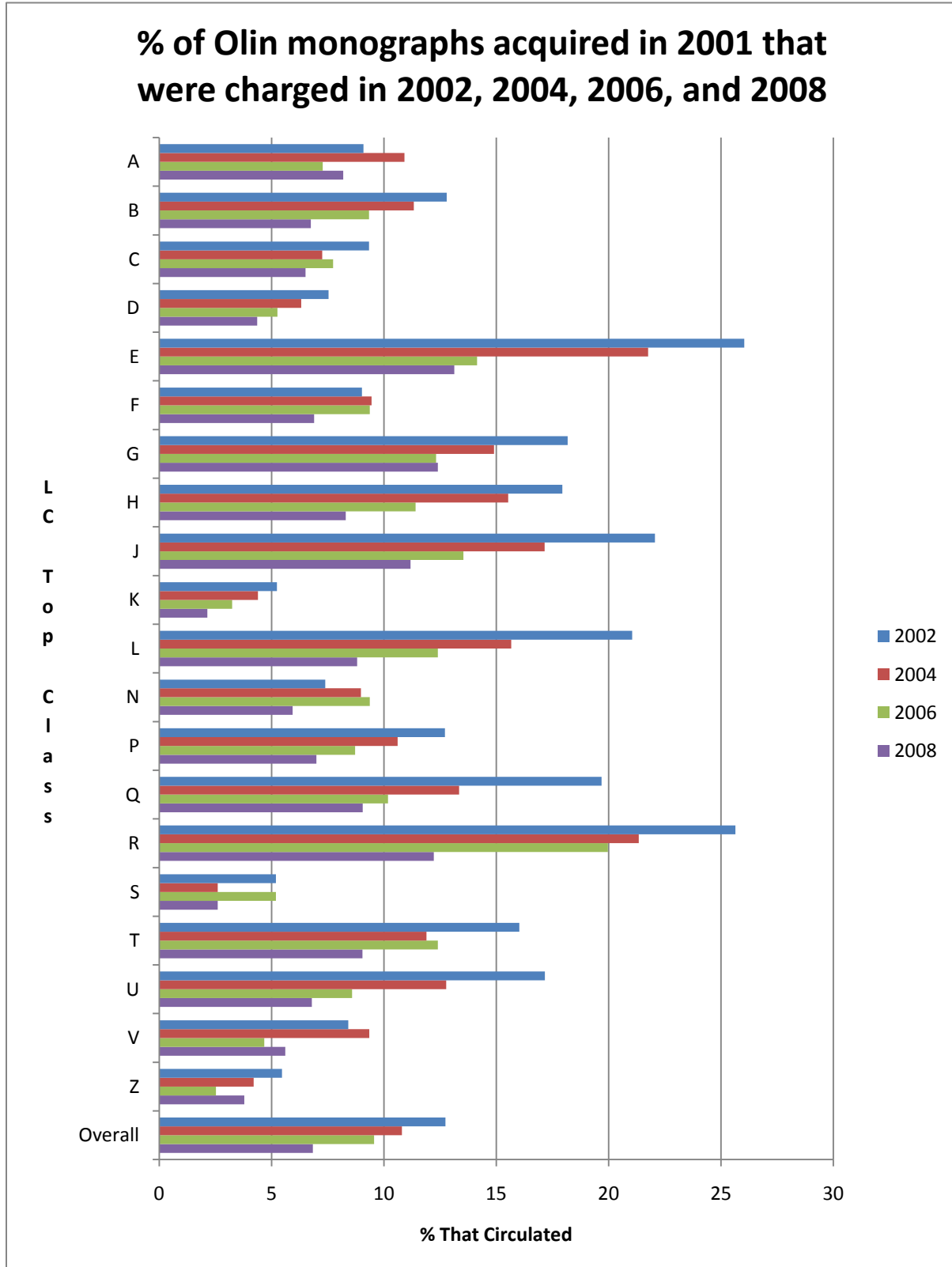


Figure 10 focuses on monographs acquired by Olin Library in 2001, breaking down the annual percentage of circulating monographs by major LC class.⁹ Olin acquired 60,469 monographic volumes in 2001, distributed among the top-level LC classes as follows:

A	110	F	2,583	L	613	R	581	Z	238
B	8,545	G	1,282	M	2	S	77		
C	814	H	7,076	N	758	T	387		
D	12,589	J	1,913	P	18,467	U	501		
E	1,379	K	2,005	Q	442	V	107		

Circulation of these volumes was highest in 2002, the first full year the entire 2001 cohort had to circulate. Certain LC classes, notably E (History of the Americas) and R (Medicine) stand out as particularly high-use, although it should be noted that the LC ranges that show the highest circulation represent relatively moderate acquisition numbers. The four top LC classes for which more than 20% of books circulated in 2002 had volume counts between 500 and 2,000. The LC classes with the largest numbers of volumes (over 7,000) in the 2001 Olin cohort had circulation rates of between 7% and 17% in 2002.

The bars for 2002, 2004, 2006, and 2008 show a fairly steady drop in circulation rates over time, at least for the higher-circulating LC ranges. By 2008, when the set was eight years old, circulation rates are more evenly balanced among the LC classes, with books in most classes circulating between 6% and 12% in that year. This suggests a strong association with subject-specific selection practices. If monographs in Medicine are acquired mainly by demand, these books are likely to be used immediately upon acquisition; their circulation would then level off, presumably.

The figure 10 graph measures percentages of Olin's 2001 acquisitions that circulated year to year, but they do not tell us about unique charges. For example, the 18% of books in the H ranges that were charged in 2002 might or might not overlap significantly with the 12% charged in 2006. Figure 11 graphs *unique* item circulation for the 2001 acquisitions from year to year through 2009 for the various CUL unit libraries. This view of the data tallies each time a volume within a set circulates for the first time. The circulation totals for each year are cumulative, including tallies for volumes that circulated in previous years. The line graph shows the growth in unique volume circulation for each unit library collection, making plain that the unit libraries differ significantly with regard to both the rate at which previously non-circulating titles continue to be checked out over time, and the point at which use of new volumes starts to level off. The average unique circulation for 2001 holdings CUL-wide was 20% by the end of 2002, 30% by the end of 2005, and 35.5% by end of 2009.¹⁰

⁹ Note that the data for 2001 acquisitions (shown in figures 10-12) does not include circulation of materials acquired that year that are housed in the Annex.

¹⁰ The charts in figures 11 and 12 do not represent the landscape as it actually was in the time range indicated (2001-2009), since they do not specify the circulation metrics from the Physical Sciences Library (PSL) that was open during this period. At the end of 2009, the highest circulating books from PSL were transferred primarily to three other libraries: 6,277 to Math, 9,670 to Engineering, and 3,247 to Mann. The highest circulating books went to the Math Library (average circulation: 6.14 times per item between 2000 and 2008). Since Math is the smallest of the three collections and received the highest-circulating books, the move had a significant impact on overall Math circulation. All PSL records, including circulation histories, were changed to the receiving libraries and it is impossible to capture these location changes. Also, our tally of charges of monographs designated in Voyager as "circulating" does not take the more restrictive circulation policies at the Law Library into account. The circulation figures we have for the Law Library, as reflected in figure 11, are therefore artificially low. Local practices at other units may also affect measurable circulation percentages.

Figure 11.

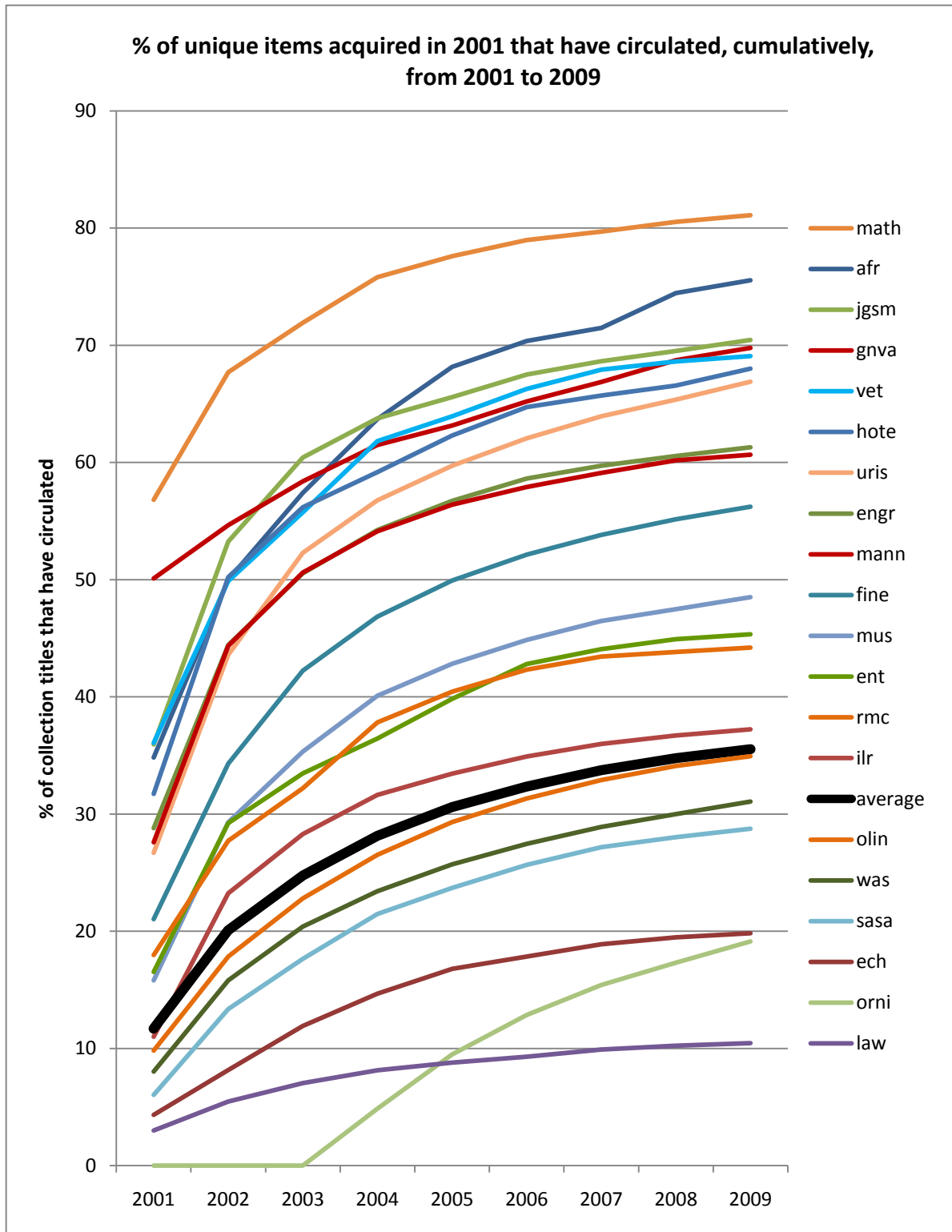


Figure 12.

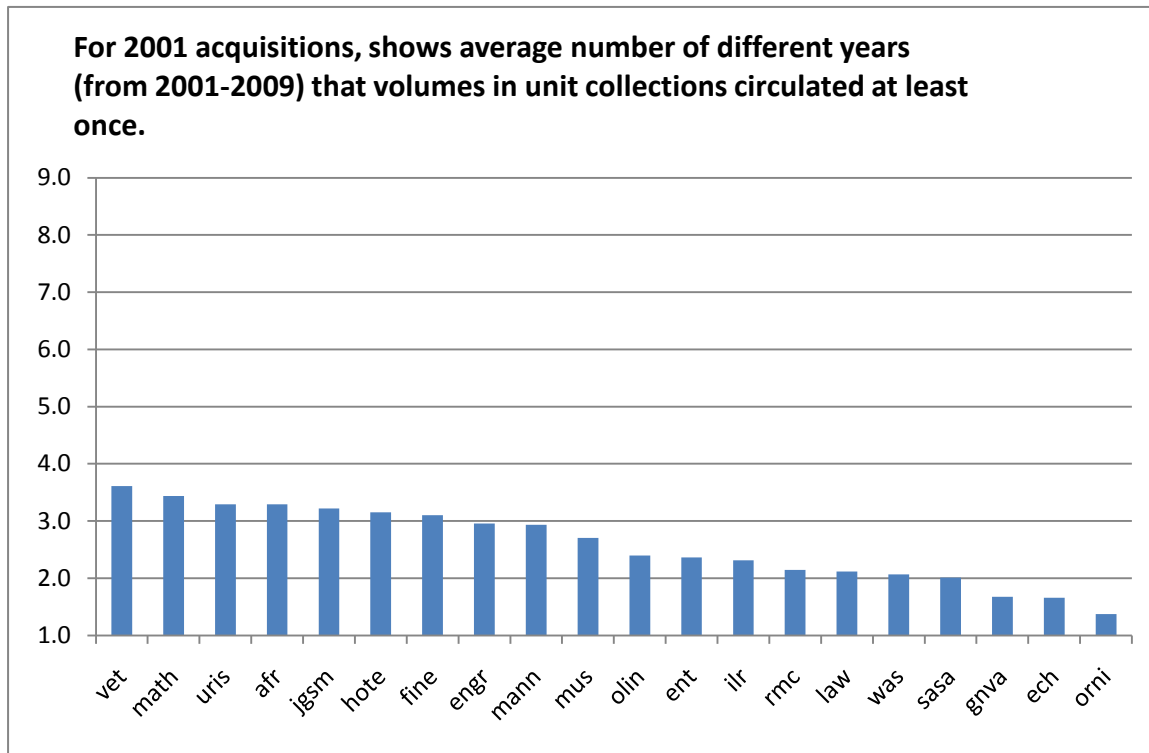


Figure 12 provides another measure of the differences among the CUL units with regard to monograph usage. The bar graph shows the number of different years in which the average volume in a unit library collection circulated, for 2001 acquisitions, over the period from 2001-2009. In this schema, the maximum value would be nine, if every volume in a collection circulated at least once in every year between 2001 and 2009. The lowest possible value is one, for a collection where every circulating volume circulated once and only once over the nine year period. (Completely non-circulating volumes are not included in this analysis). The higher the “value” for a collection in this graph, the greater the tendency for repeat circulation of volumes over a period of years. Repeat circulation can be taken as a partial measure of the continued popularity of volumes in a collection over time – their “staying power.” It should be noted that the measure is an average: it does not distinguish between a collection that might contain a core subset of books that circulate frequently and a collection with a higher number of volumes that circulate less often

Use and user communities

The historical circulation data can inform us about the use of library materials by subject (from the top LC classes treated above down to finely-grained subclasses or, ultimately, individual titles), by language, by unit library, and other data points in the title records. Beginning in 2000, the circulation records can be sorted by patron status (undergraduate, graduate, faculty, staff) as well. All of this, however, permits only conjecture about the academic program affiliation of the users of these materials. The circulation snapshot technique can foster a better understanding of how the monograph collections support particular academic programs and user communities. The snapshot analysis is a new approach, first implemented in this manner by members of the task force. Its primary advantage over previous circulation analysis techniques is that it provides more detailed and more reliable data about the users of our collections. In particular, it

allows us to correlate the circulation of materials organized by publication date, by subject area, by language of publication, and by physical library location with user data about school and college affiliations, graduate fields of study, and departmental affiliations of faculty and staff. Other advantages include the ability to more finely differentiate other kinds of circulation, such as Borrow Direct lending from traditional interlibrary loan, charges to faculty studies from graduate carrel charges, and academic staff from non-academic staff.

Potential pitfalls of the snapshot approach include the fact that it is, in fact, a snapshot – a frozen moment in time that may or may not be representative of general usage patterns. Only by repeating the snapshots over an extended period will we be able to determine how reliable an indicator they are of trends. However, the task force has taken four snapshots over a period of six months, and these suggest that circulation of the print collection does not change radically over short periods, except at obvious times such as the beginning or end of semesters.

The charts in figures 13 and 15 through 30 are based on a circulation snapshot taken on April 19, 2010 (the full data is included in the online appendices). A comparison with the historical demographic data available for Voyager-era circulation is instructive with regard to the strengths and weaknesses of both types of report. It should be noted here that, unlike the circulation reports discussed above, the snapshot includes data on circulation of CUL's entire monograph collection, regardless of the year of publication or year of acquisition.

As figure 13 illustrates, the most books in circulation when the April 19 snapshot was taken were charged to graduate students, who accounted for 34% of the total charges (affiliates of the Graduate School and graduate professional schools combined, but not including graduate carrels). Faculty had out another 23.6%, so that well over half of the 157,034 monographs charged out, CUL-wide, at the time of the snapshot were in the hands of faculty or graduate students. Undergraduates had out only 10.7% of the books charged – 16,744 books in total or an average of about one book per undergraduate student in the Cornell population (compared to approximately 8 books, on average, for graduate students and about 13 per faculty member). The snapshot numbers taken alone are potentially misleading, however. A longer view, using the circulation records for books acquired in 2001, reveals that undergraduates charged out these books more frequently over the past decade than any other campus demographic: 33.2% of the total charges of 2001 acquisitions by mid-April 2010 (see figure 14). The dramatic difference between the snapshot and the ten-year view points to frequent turnover in undergraduate circulations, which presumably reflects different research needs (short-term, class-based projects). The short circulation period, compared to graduate students, staff, and faculty, that the Library allows undergraduates is based, in part, on a recognition of these different research needs. Note that faculty charges amount to only 11.3% of the total charges of the 2001 monograph cohort, about half of the proportion of faculty usage shown by the snapshot. Here, again, we can assume that the difference reflects intensive, long-term work in relatively circumscribed areas, compared to more extensive, exploratory use of the collection that characterizes student research. Another factor to bear in mind is the relative longevity of faculty and staff on campus compared to the relatively quick turnover in the student population: graduate and undergraduate use of the collection since 2001 represents the research of multiple student generations. The table below shows the average number of days books had been charged out and the average number of renewals of the books, by borrower status, for the April 19, 2010 snapshot. The average item charged to a faculty member had been out for 644.6 days; the small number of average renewals is reflective of the influence of indefinite loan. Books charged to Graduate School students, by comparison, had been out an average of 340 days with 2.3 renewals. Books charged to undergraduates had been out an average of 78 days, with 2.1 renewals. By way of contrast, note that the average item charged a faculty study had been out for over seven years.

Status of borrower	Average days out	Average number of renewals
Academic Staff	674.6	3.6
Borrow Direct	27.8	0.0
Carrel	424.1	0.9
Faculty	644.6	0.5
Faculty Study	2616.2	7.5
Graduate Professional	122.2	0.6
Graduate School	340.1	2.3
ILL	37.8	0.3
Internal	342.5	1.9
Other	297.1	11.9
Staff	654.3	4.1
Undergraduate	78.2	2.1

The snapshot data makes visible the departments and graduate fields that stand out as heavier and lighter users of the monograph collections. And by showing up the breadth of collection usage by different programs and pointing to LC classes that appear to have high “cross-over appeal” to multiple fields, the data usefully complicates an easy identification of the Library’s subject classifications with particular academic programs. Figures 15 through 34 are provided as examples – from the humanities and the sciences – mapping subjects to academic programs and programs to subjects. For these examples, we have chosen the following LC sub-classes:

- B Philosophy, general (not to be confused with the “B” top class, which covers philosophy, psychology, and religion)
- PR and PS English and American literature (combined here)
- QA Mathematics
- QD Chemistry

Charts for the top class S (Agriculture) are also included here. For the books in these classes that were charged to faculty or graduate students, the charts break down the charges by department or graduate field, giving the percentage of books charged to patrons in the various programs. Only the ten programs with the highest representation are named; the rest are collapsed into an “other” category. We have included charts for the following academic programs:

- Department of Philosophy
- Department of English
- Department of Mathematics
- Department of Chemistry and Chemical Biology
- Department of Horticulture
- Graduate Field of Philosophy
- Graduate Field of English Language and Literature
- Graduate Fields of Mathematics and Applied Mathematics
- Graduate Field of Chemistry and Chemical Biology
- Graduate Field of Horticultural Biology

These charts show the proportion of books charged to faculty and graduate students in the programs in the ten LC sub-classes with the highest representation.

The wide-ranging interchange among the disciplines suggested by these charts is striking, although it is also interesting to see significant differences among academic programs in the degree to which they are using books from beyond their fields. Not surprisingly, perhaps, a substantial divide between the

humanities and social sciences, on one hand, and the sciences, technology, and mathematics, on the other, is plainly visible in these examples, although there are notable instances of interpenetration. More granular analysis of the relationship between books and academic programs is possible from the snapshot data. For the time being, we offer a few preliminary observations:

Circulation of books classed as “general philosophy” was fairly widely and relatively evenly distributed among several humanities disciplines. While members of the Philosophy Department accounted for nearly a quarter of the faculty charges, five other humanities departments each had between 7% and 14% of the books charged to faculty. Among Philosophy faculty, books classed in QA (mathematics) represented the third largest number of books charged, after books in the B sub-class and in PA (Classics). Books in mathematics do not show up in the top ten LC sub-classes for graduate students in Philosophy.

Books in English and American literature did not travel as far afield: 55% of faculty charges were to members of the English Departments and 63% of the charges to graduate students were to students in that field. Nearly half of the books charged to graduate students in English were in the PR and PS ranges, though these made up only 30% of charges to English faculty.

Of the books in mathematics (QA) charged to faculty, a full 43% were to members of the Mathematics Department (compared to 19% of graduate students in Math for the graduate charges); the rest of the top ten faculty charges were in rather wide-ranging fields, including Economics and History. Among the mathematics books charged to graduate students, the top ten fields did not include any from the humanities or social sciences. Books charged to faculty in Mathematics and to graduate students in Mathematics and Applied Mathematics were overwhelmingly in the QA sub-class (74% for faculty, 70% for graduate students), but each patron group also had books charged in a variety of subject areas, including several in the arts and humanities.

Chemistry (QD) books charged to faculty were likewise largely out to faculty in Chemistry (41%), but in this case a sizable portion were charged to humanities faculty (9% to faculty in Romance Studies, 4% to faculty in History). Graduate students in Science and Technology Studies had out 3% of the graduate charges of the QD books; the rest of the top ten fields were in the sciences in the strict sense. The faculty and graduate student charges from the Department and Graduate Field of Chemistry and Chemical Biology included significant numbers of books in the arts and humanities.

Faculty charges of books in Agriculture (LC top class S) were widely distributed among departments – primarily in the College of Agriculture and Life Sciences, but it should be noted that 12% of the faculty charges were to members of the Department of Architecture. No single department had out more than 14% of these books that were charged to faculty. Graduate student charges of this material were likewise widely distributed and, beyond the life sciences, included significant circulation to graduate students in City and Regional Planning and in History.

Of course, we cannot distinguish between professional and leisure or avocational use of library materials. A possible explanation for the wider range of subjects charged to faculty, compared to graduate students, in particular areas of study, is simply that faculty have more time to pursue outside interests.

Figure 13.

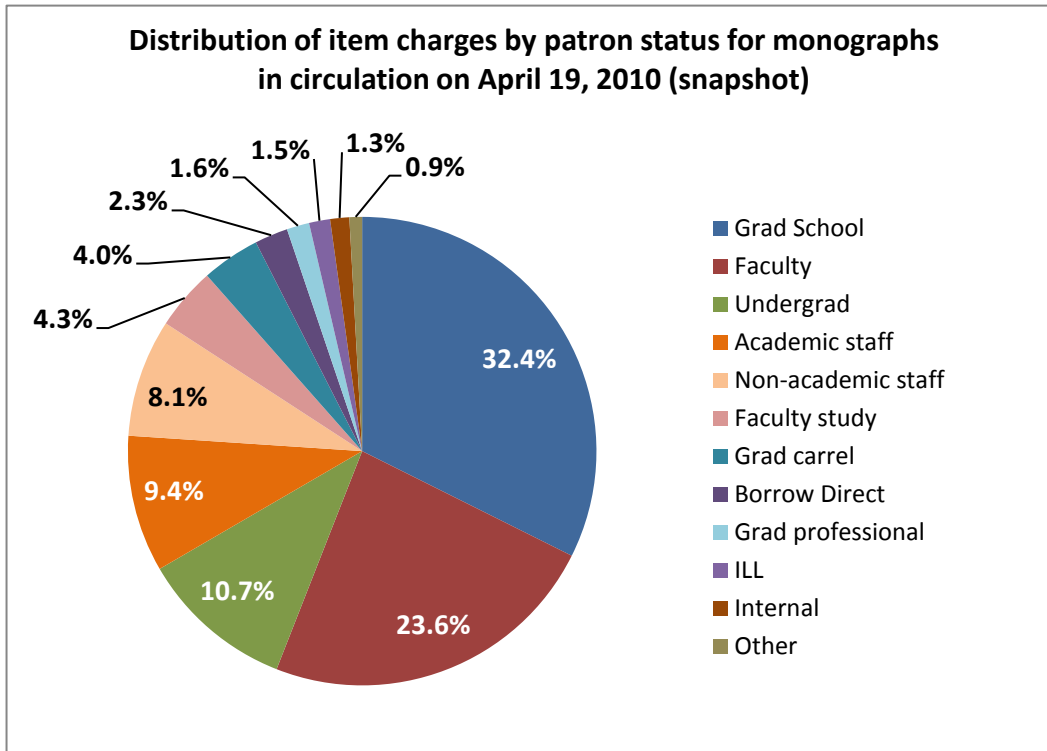


Figure 14.

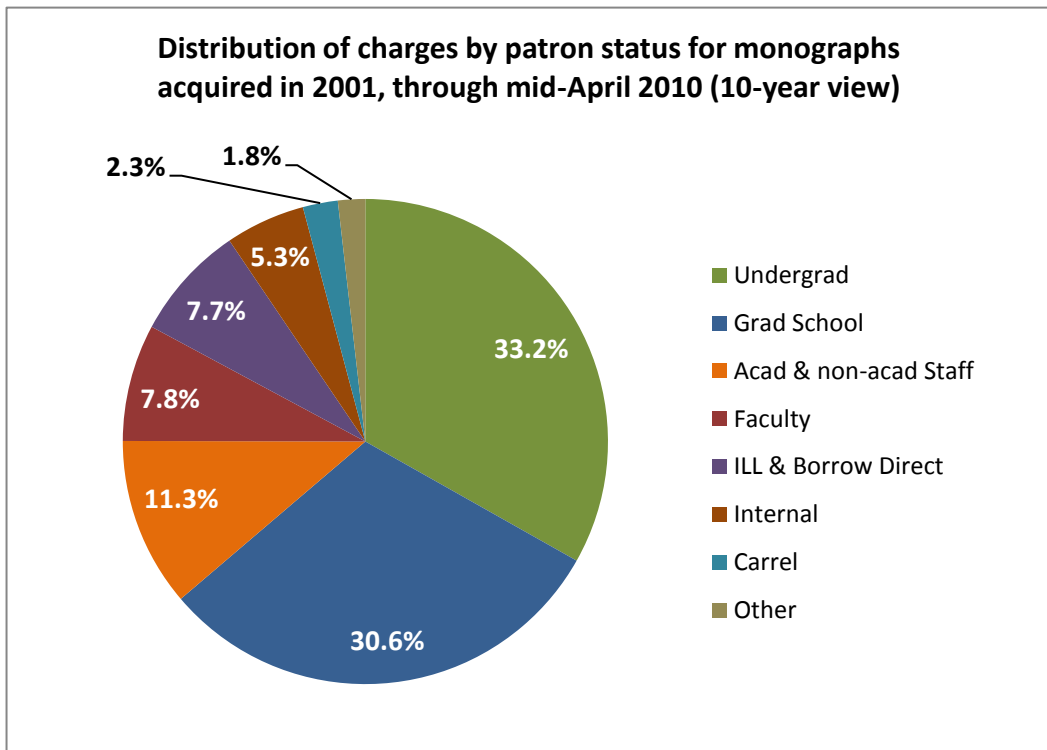


Figure 15.

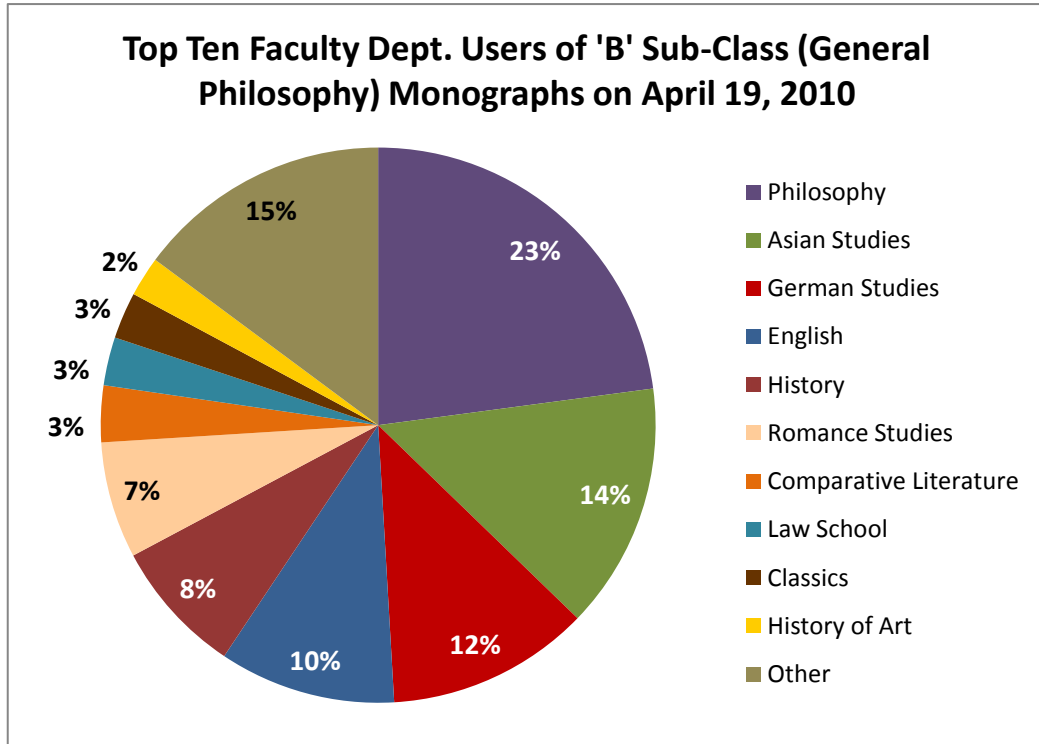


Figure 16.

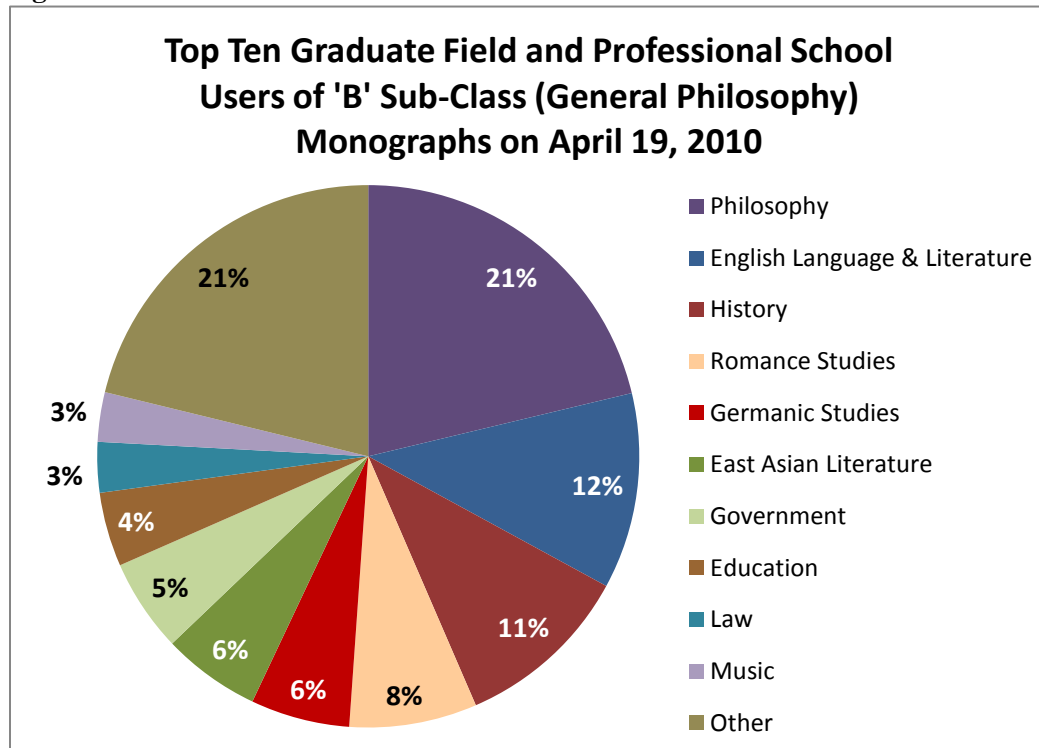


Figure 17.

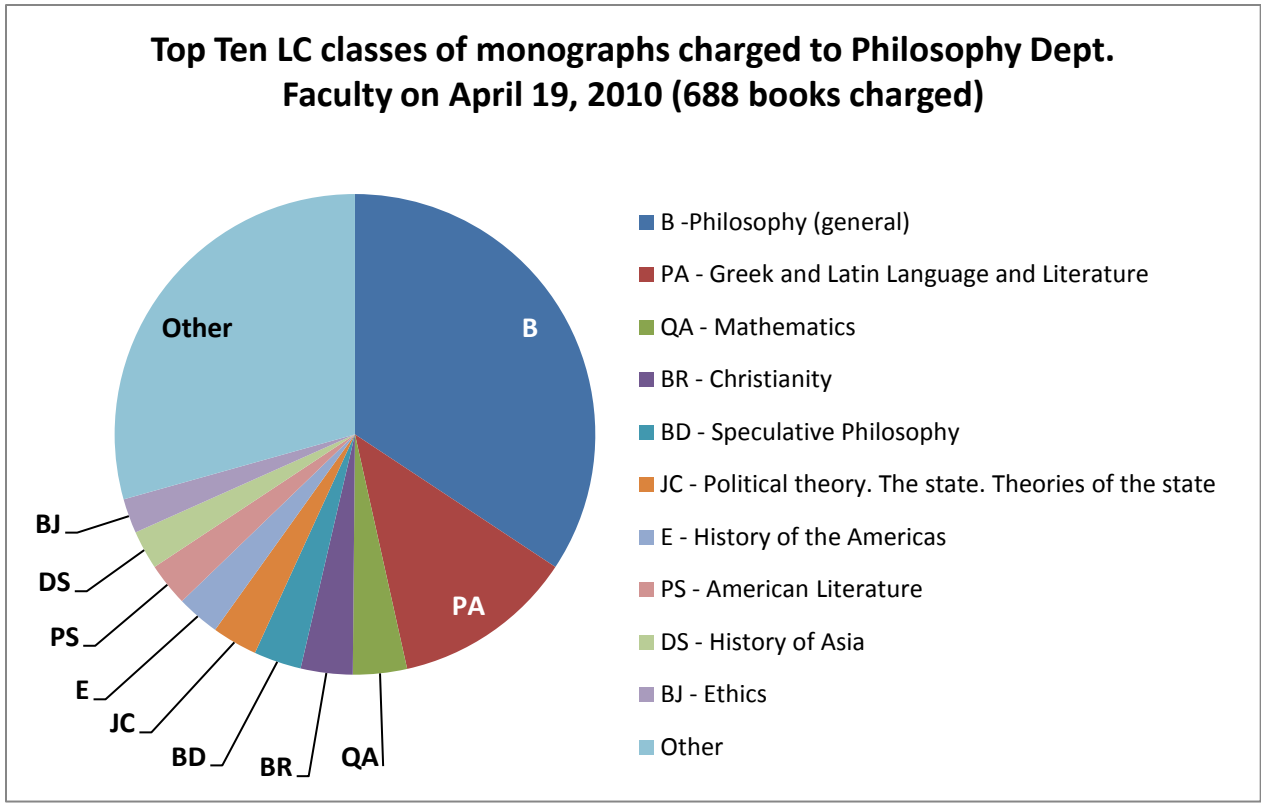


Figure 18.

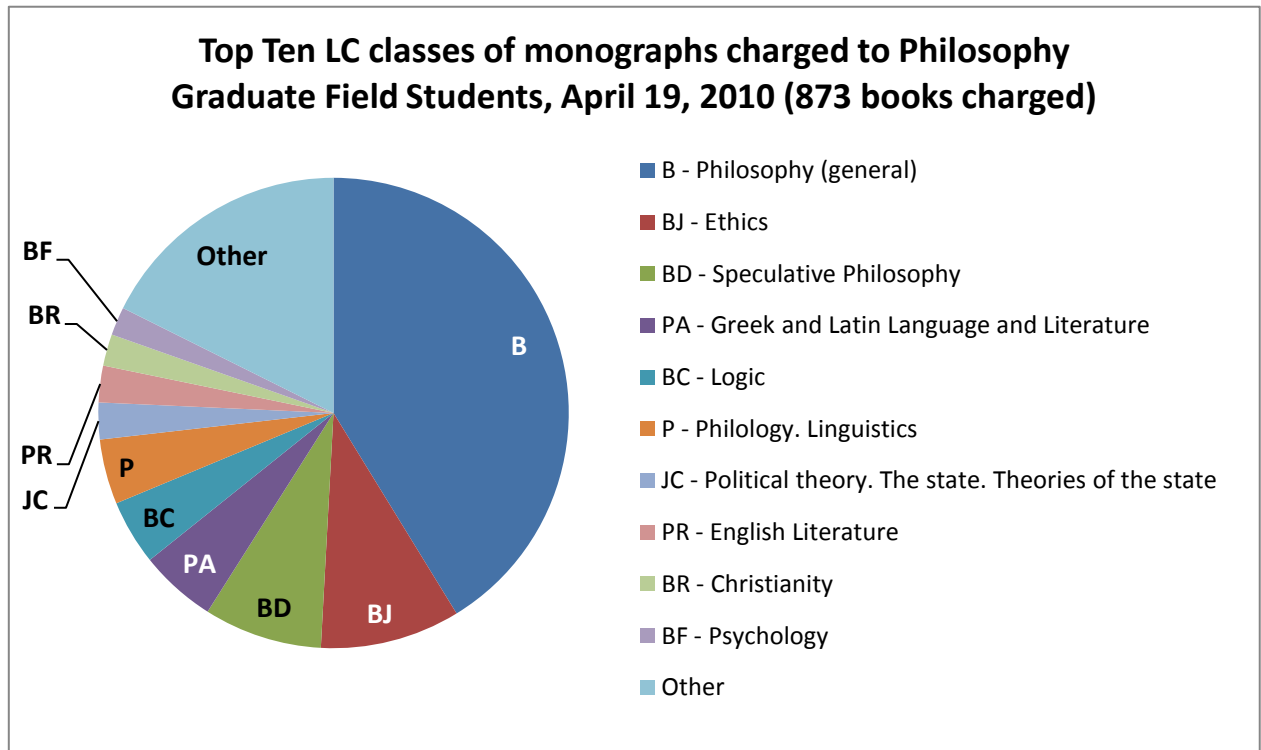


Figure 19.

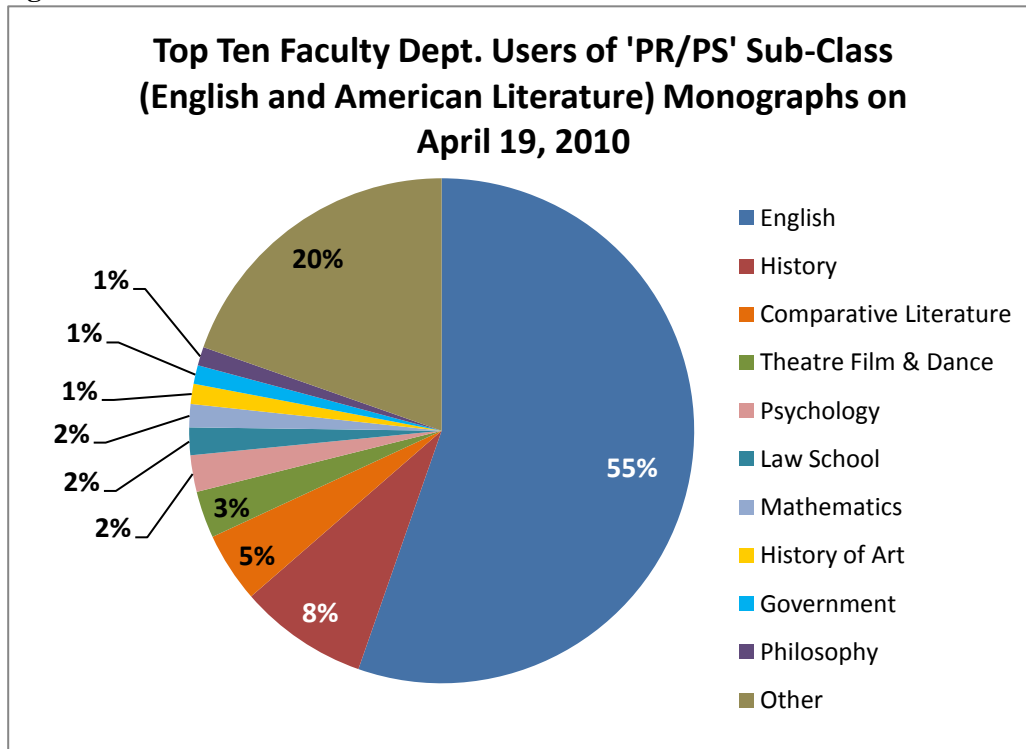


Figure 20.

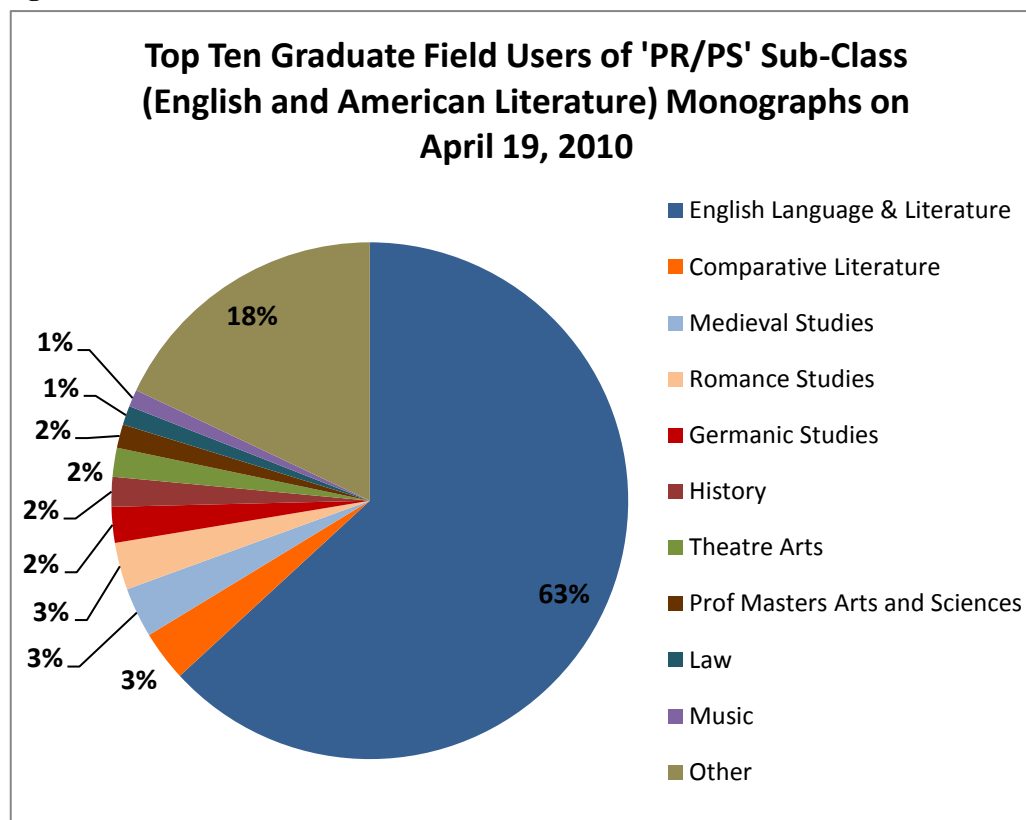


Figure 21.

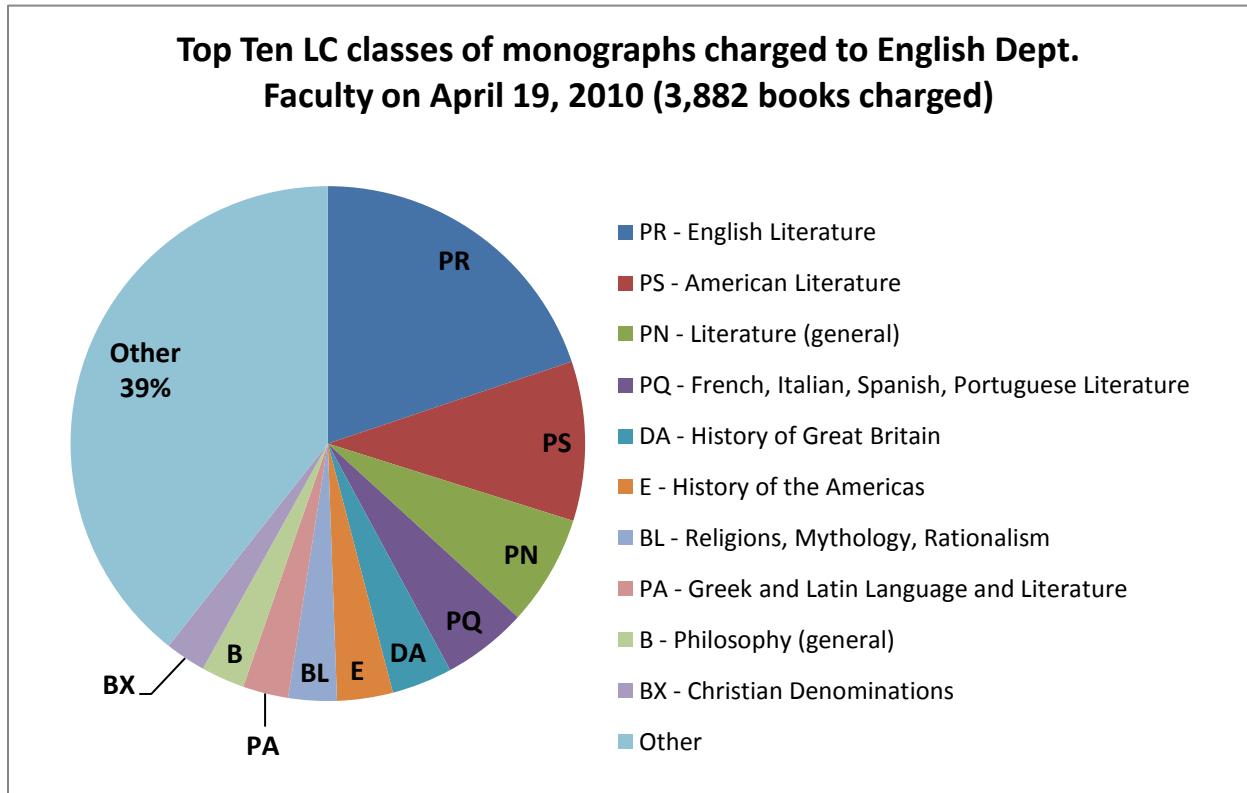


Figure 22.

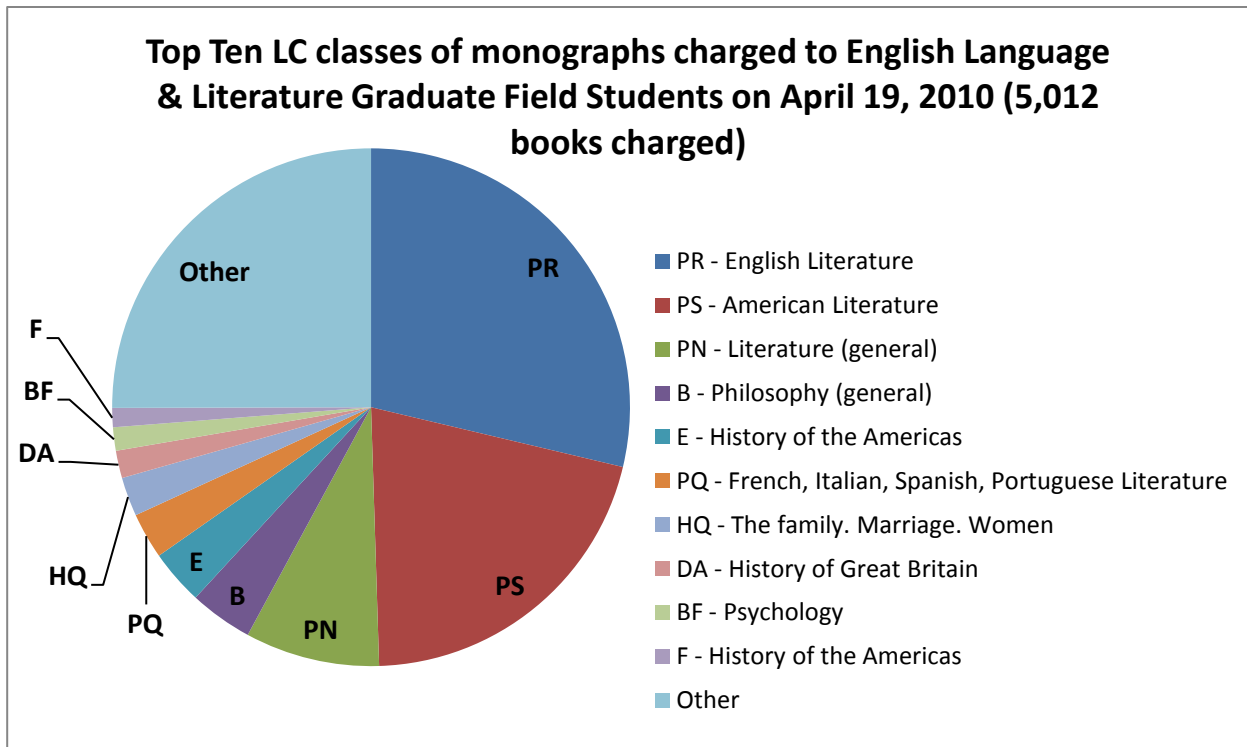


Figure 23.

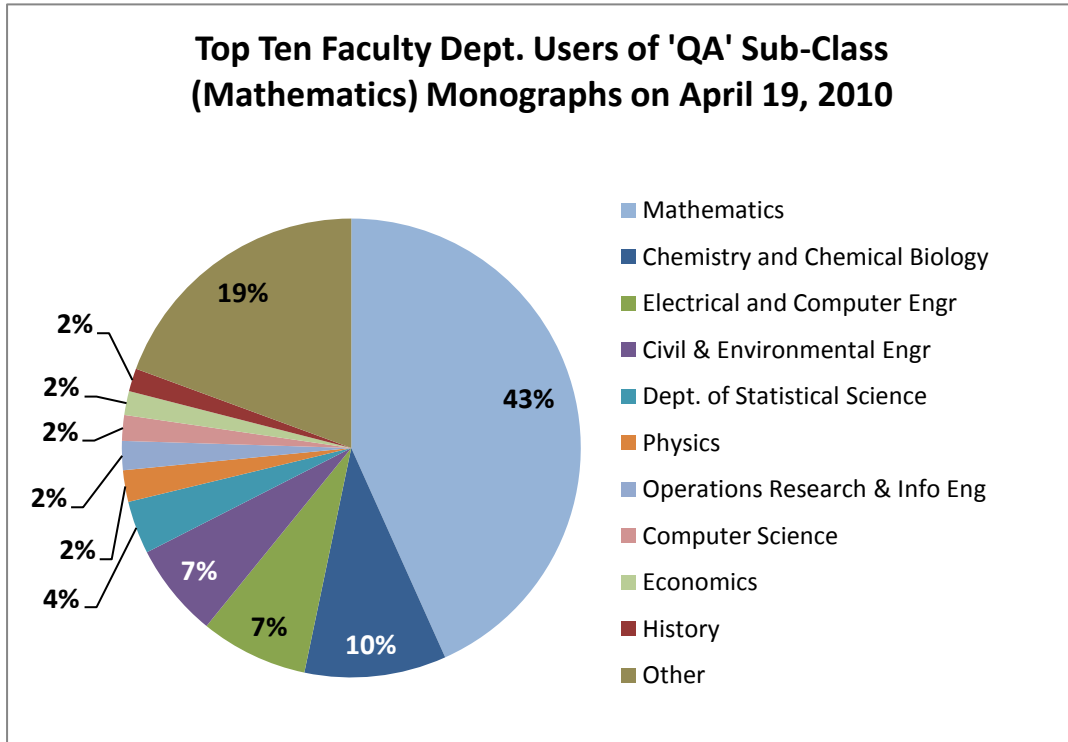


Figure 24.

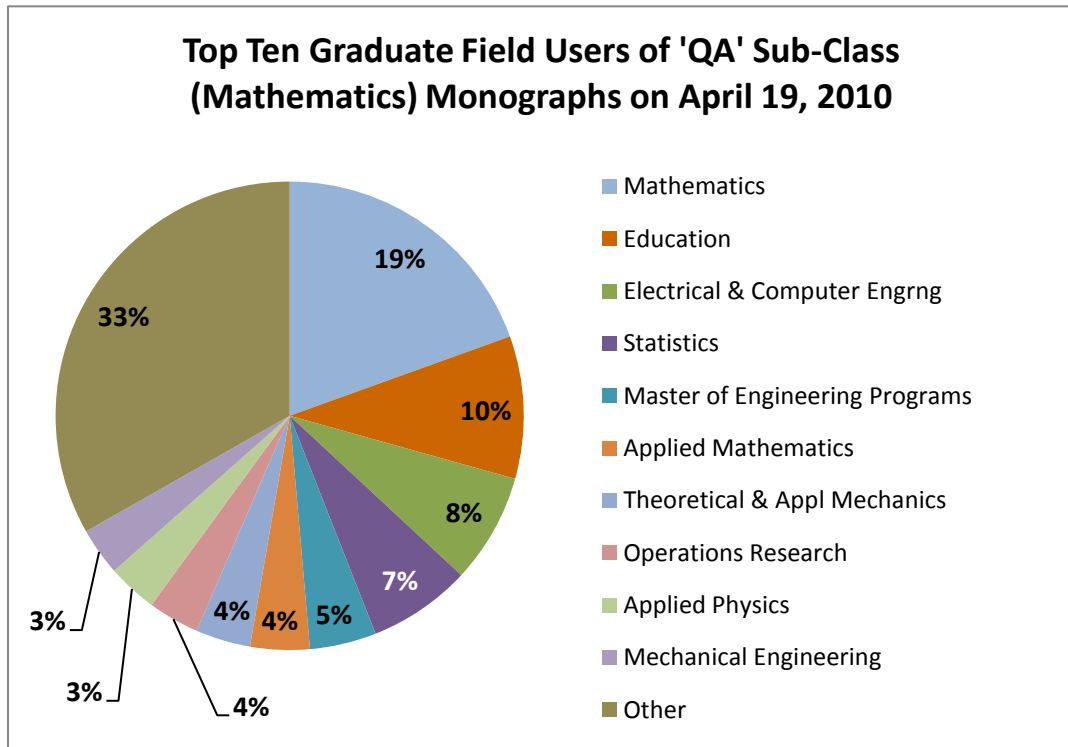


Figure 25.

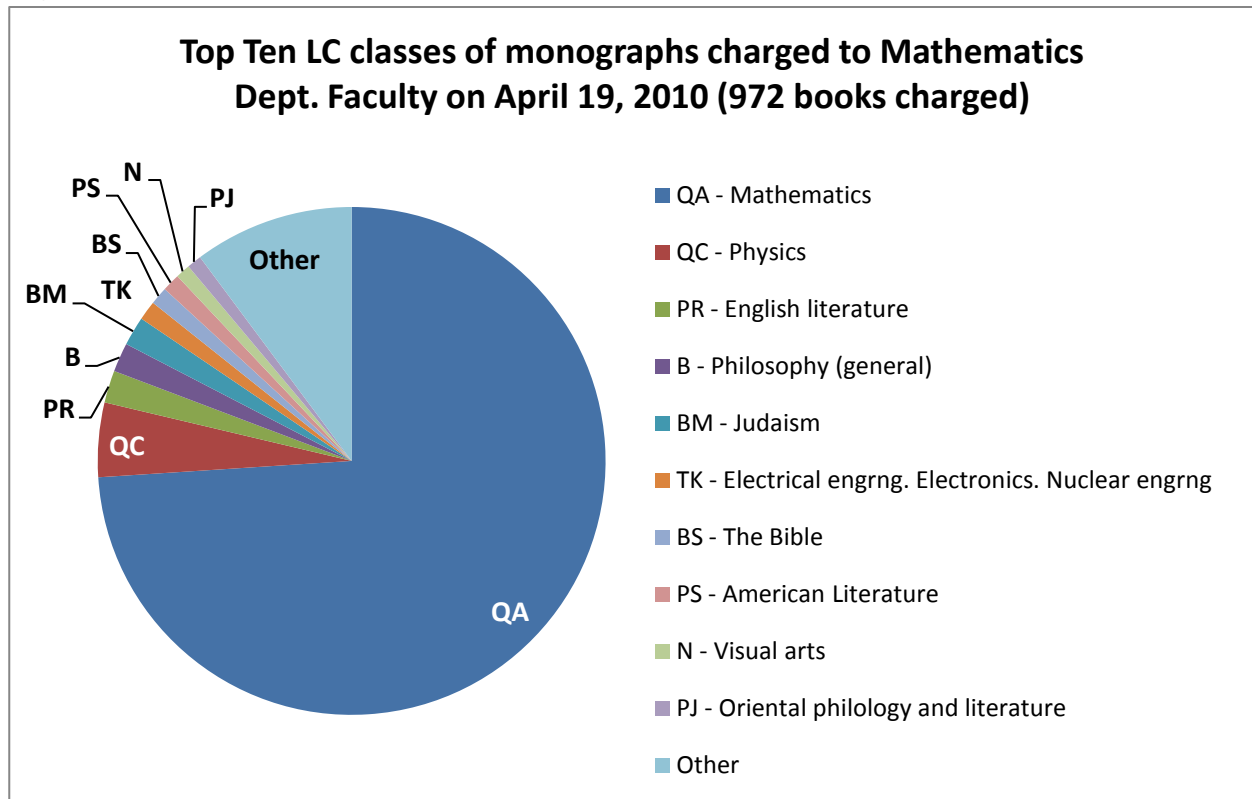


Figure 26.

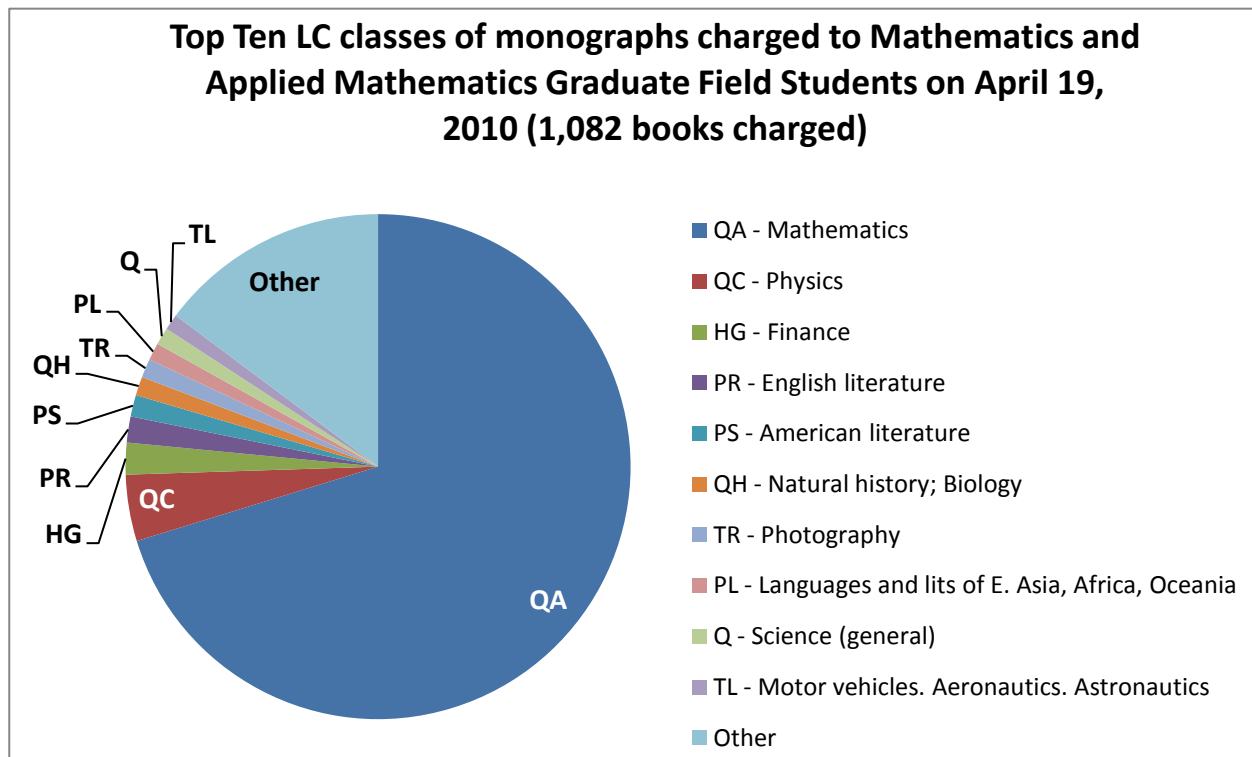


Figure 27.

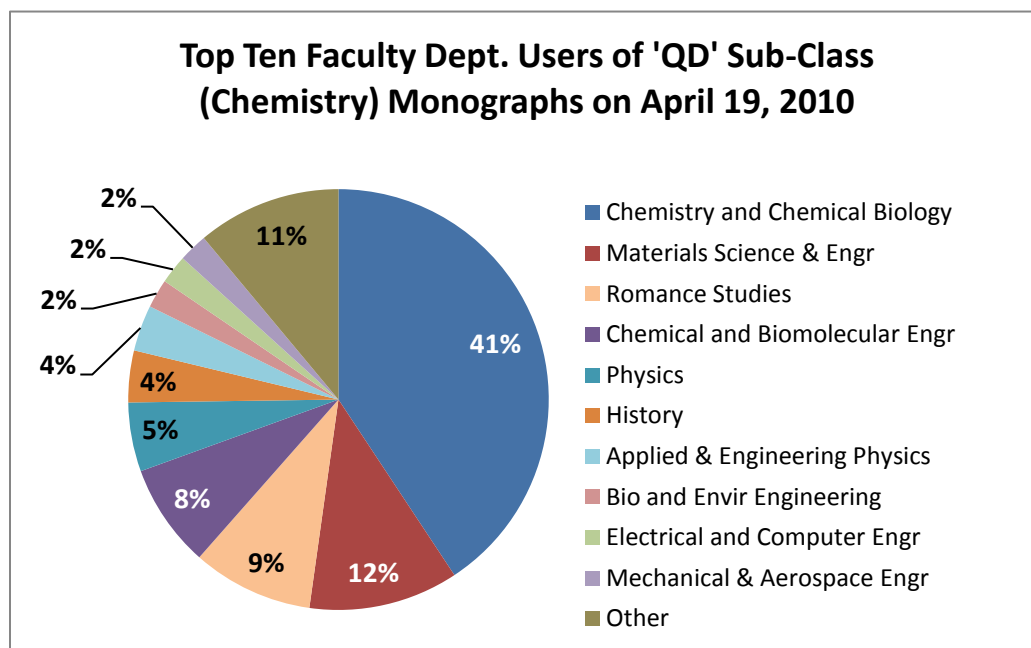


Figure 28.

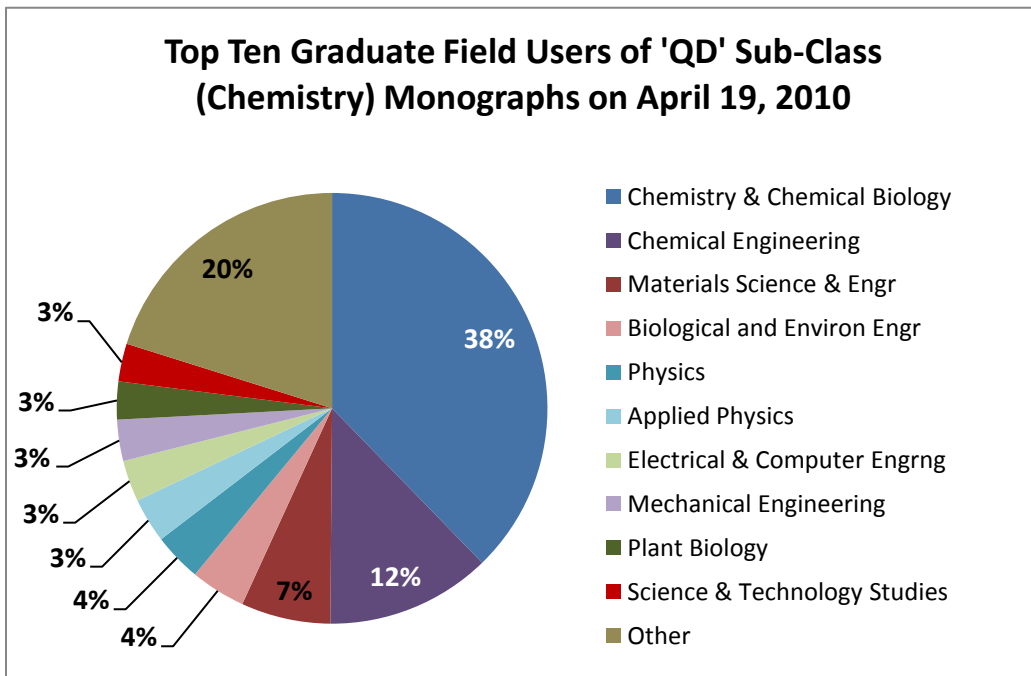


Figure 29.

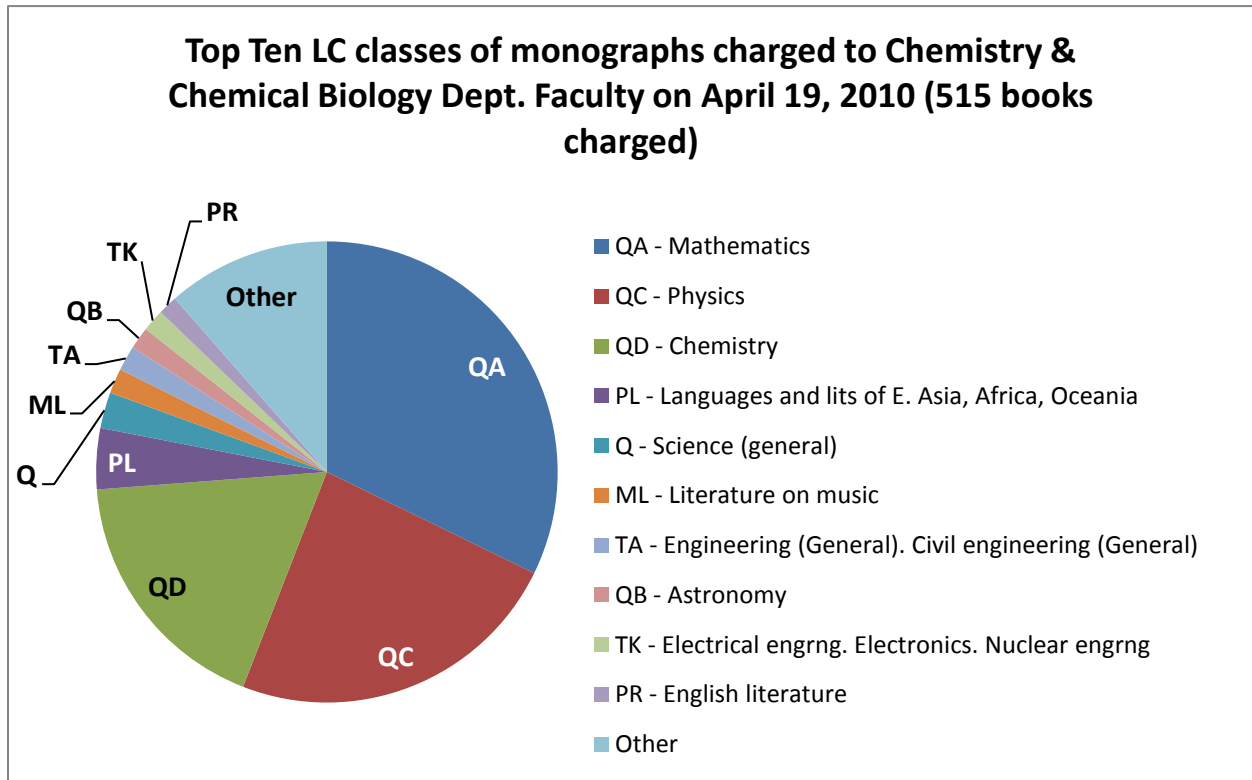


Figure 30.

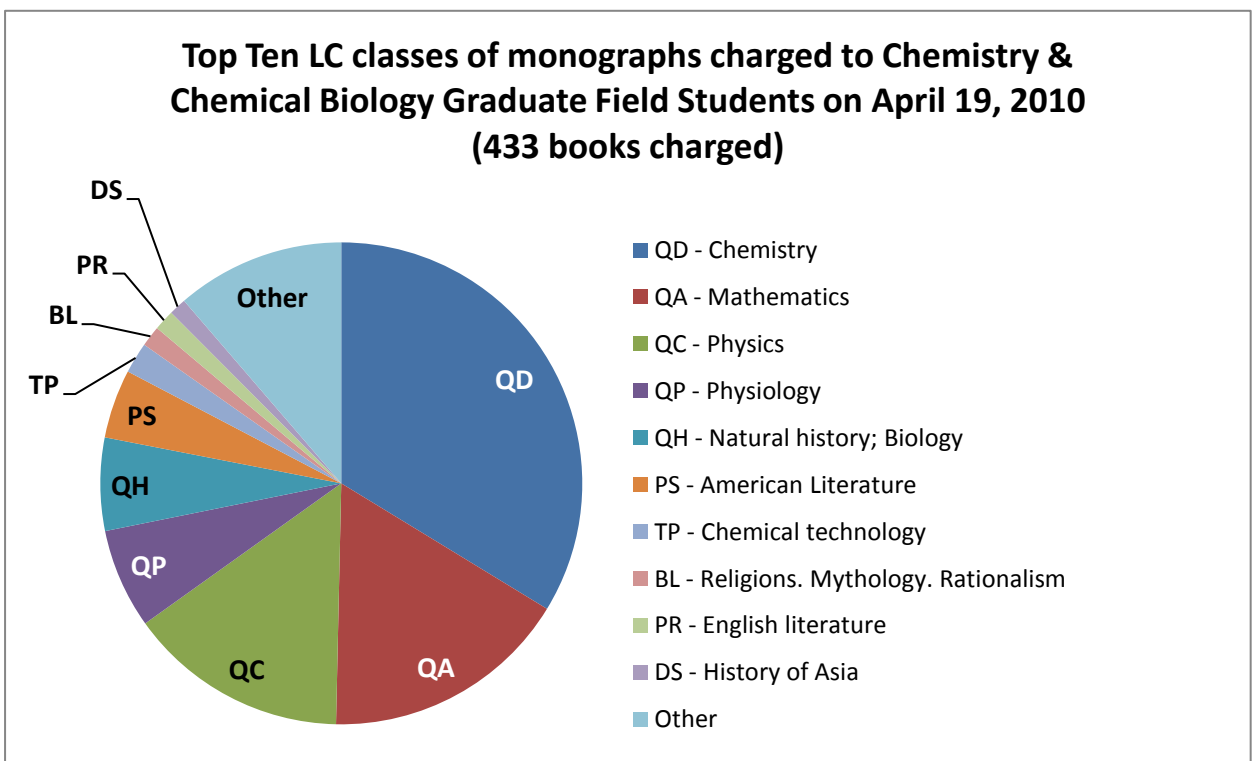


Figure 31.

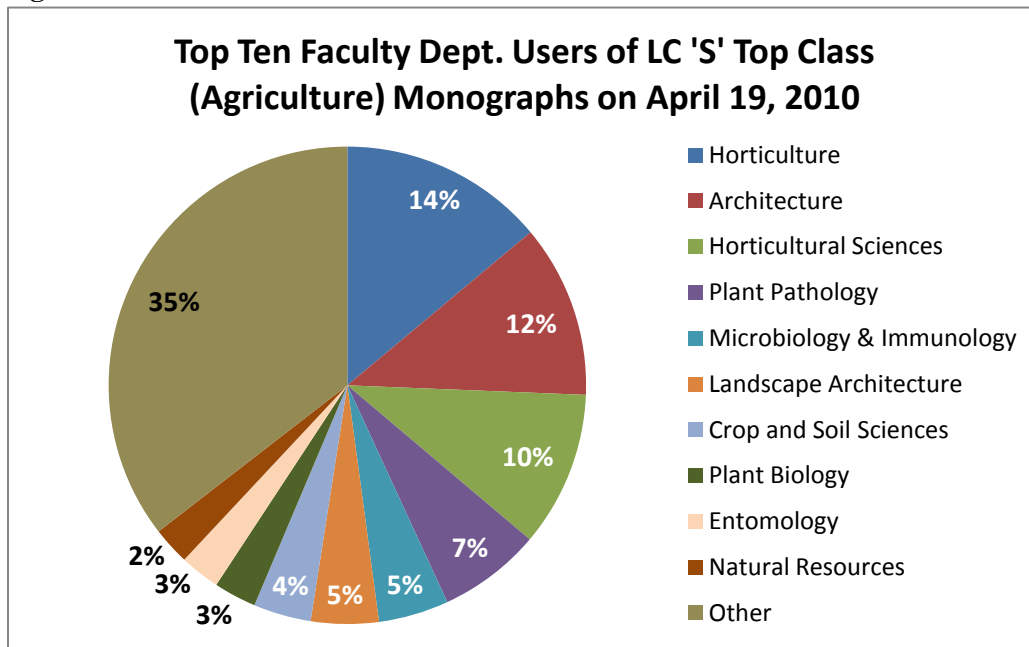


Figure 32.

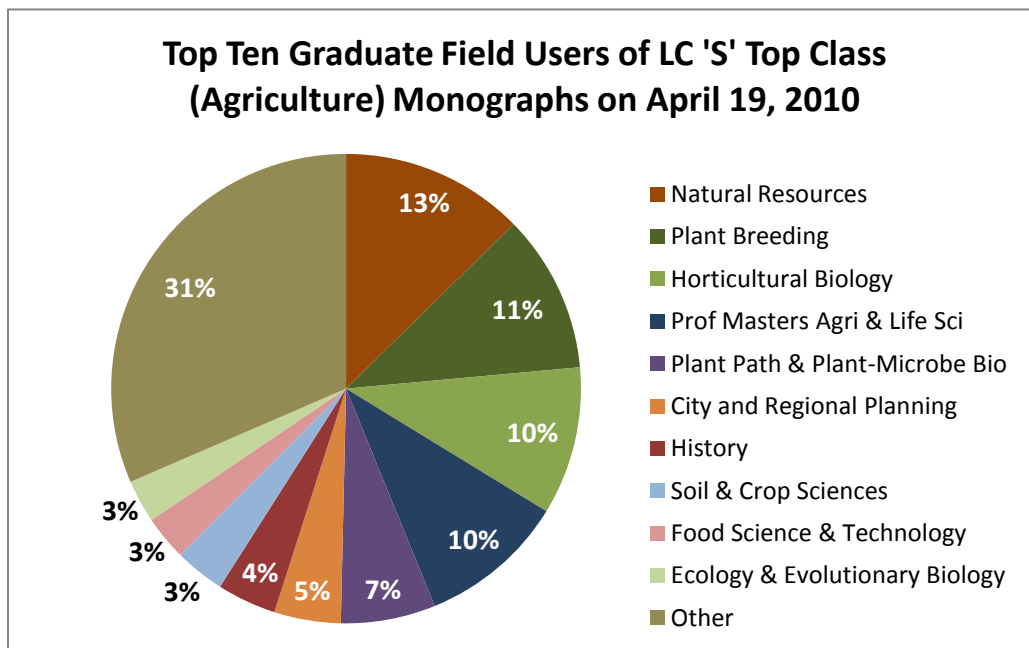


Figure 33.

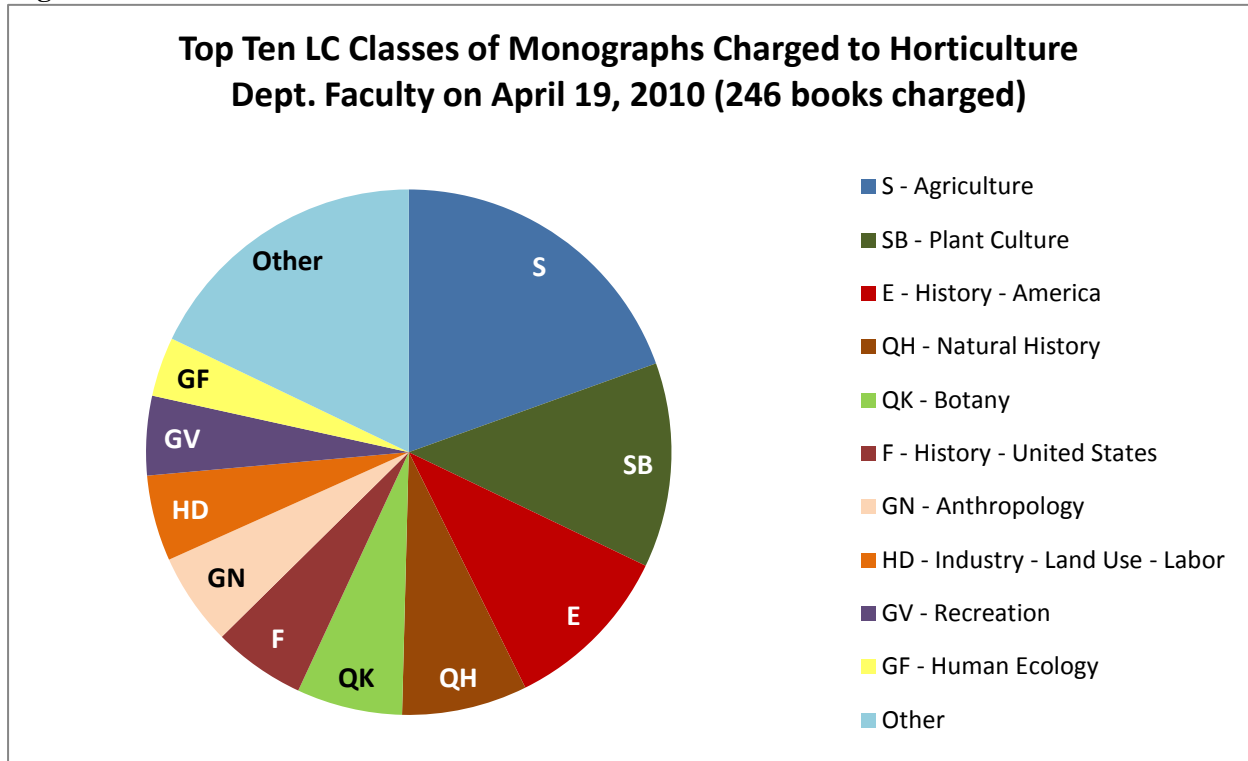
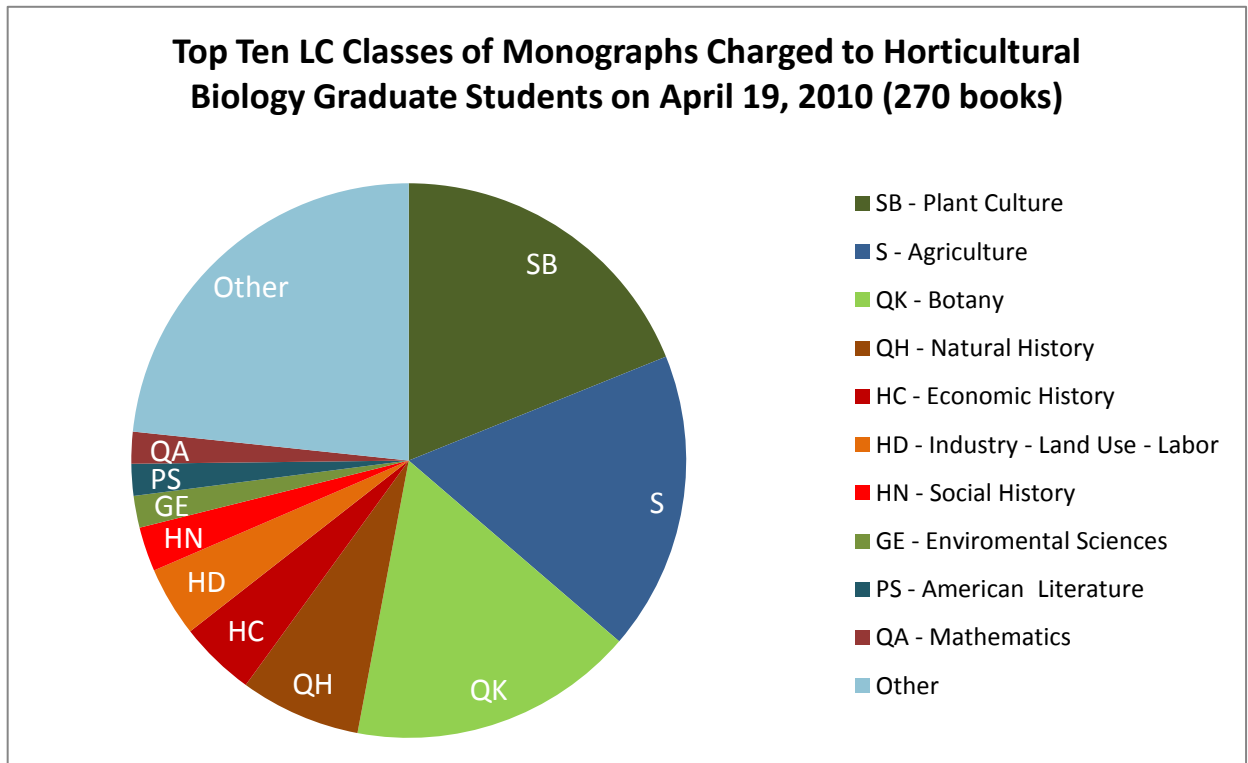


Figure 34.



Conclusions and Recommendations

Considerations of monograph usage have long driven collection development decisions in public and smaller academic libraries where limited space, personnel and money necessitated such a course. Monographic collections in the largest research libraries developed along a different route, with a preference towards increasing space and budgets as needed over time in order to allow for the purchase of as many monographs as possible. The main limiting factor at these libraries was whether or not a particular monograph fell within the scope of the collection. Scope, in most cases, was determined over the years by the research and teaching interests of the institution's faculty and students. Thus, millions of books were collected and libraries judged mainly by those numbers.

Problems encountered with space and budgets towards the end of the 20th century and increasingly over the first decade of the 21st century have caused even the largest libraries to reevaluate their collection development programs. The philosophy of purchasing monographs "just in case" they may be needed at some point in the future by unidentified researchers has come under particular scrutiny, also at Cornell. Even before the current economic crisis, Library budget realities in the face of rising costs dictated, at least for certain fields, that CUL purchase monographs primarily in response to immediate demand. As e-books emerge as a viable format for providing access to monographic content and various patron-driven acquisitions models mature, CUL will no doubt move further in the direction of "just-in-time" collection building. Improved interlibrary borrowing networks and the potential for meaningful collaboration in collection building among institutions can also be expected to shift some emphasis away from self-sufficient, locally-built collections. What can the data about print monograph circulation tell us about the balance of prospective collection development and demand-driven purchasing that is appropriate to the Cornell context?

Our study confirms that a significant portion of the CUL stacks is taken up by books that have never circulated over a period of decades. What significance the Library and the University should assign to this non-circulating material is far from clear, however. If half of CUL's monograph purchases of the past twenty years have circulated, is that a lot or a little? Precious resources are being spent to purchase, house, and preserve these books, but to what extent should this be regarded as misspent funds and to what extent as investment in a strategic reserve? The answer will surely vary by field and by the intended readership for particular segments of the collection. Factors such as language of publication can place distinct limits on the pool of potential users and any meaningful measure of usage must take the size of the user population into account.

Our study shows that the print monograph collection is used extensively and suggests that its use today is comparable to circulation over at least the past two decades (although some of this has to remain conjecture, due to the limitations of the older data, as we have discussed). Use is widely distributed among undergraduate and graduate students, faculty and staff, and books in the various subject areas are used by patrons in a multiplicity of academic fields.

Patterns of use that can be detected at a high level of abstraction (the level at which we have analyzed usage in this report) resolve to a highly complex interplay of old and new subject matter, fluctuating fields of study, interdisciplinary cross references, expert knowledge and new discovery (among many other factors), when viewed at a more granular level.

The task force makes the following recommendations:

1. Integrate monograph usage data and its assessment into CUL's collection development operations at the administrative level.

CUL's Scholarly Resources and Special Collections division should oversee the regular collection and evaluation of usage data on the monograph collection. This should include a routine schedule for performing circulation snapshots; periodic snapshots will allow the development of a repository of data that can be used to track changes in the circulation pattern of materials over time. Better historical statistics will allow us to track factors such as the impact of increasing e-book availability on the use of print in particular fields and among particular user groups. We recommend that the impact of Annex transfers on the use of print also be traced. Ongoing, systematic analysis of ILL and Borrow Direct statistics should complement the evaluation of local use of the local collection. CUL should monitor work in collection usage assessment at other at other institutions and keep up with best practices.

Except for the snapshot data, the present study only looks at circulation of books acquired since 1990. It would be instructive to compare this data to the recent (Voyager-era) circulation of CUL holdings as a whole. The data for pre-1990 publication dates could not be assembled in time for this report, but it will be available at a later date, and could be the basis for a follow-up analysis.

2. Make past monograph usage, understood in context, one facet in decision making about future CUL acquisitions and investment in the collection.

Because any interpretation of usage must proceed from an understanding of specific disciplines and patron groups at Cornell, we recommend empowering the subject librarians who select materials for CUL's collection with data about the use of the parts of the collection that they oversee. Consideration of monograph usage should be part of the materials budget allocation process, but this cannot simply mean that low-use parts of the collection are defunded. Rather, selectors should be able to speak knowledgably, in the context of the budget allocation process, about what the usage numbers mean for the collection they are trying to build and why this is important for scholarship at Cornell. Considerable further discussion in CDExec and CUL administration more broadly is clearly needed to refine expectations surrounding monograph usage and collection development. The task force recommends firmly against the adoption of specific across-the-board targets for the circulation rate of print monographs acquired for the collection.

3. Make monograph usage data accessible to subject selectors.

Empowering selectors to interpret monograph usage data for the disciplines they serve requires that monograph usage be addressed in selector training and continuing education, and also that tools be developed that will facilitate usable views of the circulation data. Selectors now have access to various predefined "live" reports on materials budget expenditures, and we should add customized, dynamically-generated usage reports to the selector toolkit. Selectors need rich, easily accessible information about monograph usage to inform selection decisions as well as collection assessment activities.

4. Let a more detailed understanding of usage inform where and how we house books.

To date, CUL has moved a large amount of material to the Annex by relying on circulation statistics. Monographs published before a certain date (the date differs by collection) with no circulations have routinely been transferred to the Annex. Analysis of the usage data might allow the Library to target

certain groups of materials much earlier in their life. Rather than leaving a non-circulating book in the stacks for twenty years before transferring it to the Annex, the data could be used to tell us how likely that book is to circulate based on patterns of usage for similar items. Caution should be exercised here to make certain that other factors affecting usage are first mitigated, such as the difficulty of finding books written in non-Latin based scripts in the catalog.

5. Consider how usage data for the academic programs relates to the size of these units and Library investment in the collections that support these fields.

We have access to demographic data about faculty, staff, and students in Cornell's schools and colleges, departments, undergraduate majors, graduate fields, etc. In order to fully evaluate the circulation data, we need to understand the size of these patron groups and how this relates to collection usage. Such an analysis would permit a better understanding of whether the Library is adequately and equitably investing in the research needs of the various academic programs. This analysis would need to extend well beyond print monographs to take the journal literature, as well as electronic formats, into account.

We need to explore whether it is viable to associate investment in materials with particular academic programs given the increasingly interdisciplinary nature of much scholarship.

Appendix 1. Print Collection Usage Data on the Task Force wiki

The following data reports are linked from the the Collection Use Study Task Force page on CUL's Confluence wiki: <https://confluence.cornell.edu/display/culcdsc/CUL+collection+use+study+%282009-10%29>.

- I) Holdings data
 - A) Holdings by LC class per unit
 - 1) all monographs
 - Counts of all monographs by unit library and LC class. Each unit library on a separate worksheet tab. Music scores on a separate tab, with all locations on a single tab. Includes acquisitions through mid-April 2010 for items published 1990 or later, and through late June for items published before 1990. Excludes cons and cts locations. Maps grouped with Olin. Sublocations included with main location, except for annex locations, which are shown separately on the same tab as the associated main location. Excludes lost/missing/withdrawn items.
 - 2) monographs published 1990-2010
 - Same as above, except covers only monographs published 1990-2010. Thus it is a subset of I.A.1.
 - B) Holdings by top level LC class and language, per unit
 - 1) all monographs
 - Counts of all monographs by unit library, LC top class (first letter of LC class), and then language. Each unit library on a separate worksheet tab. Music scores on a separate tab, with all locations on a single tab. Includes acquisitions through mid-April 2010 for items published 1990 or later, and through late June for items published before 1990. Excludes cons and cts locations. Maps grouped with Olin. Sublocations included with main location, except for annex locations, which are shown separately on the same tab as the associated main location. Excludes lost/missing/withdrawn items.
 - 2) monographs published 1990-2010
 - Same as above, except covers only monographs published 1990-2010. Thus it is a subset of I.B.1.
 - C) Holdings by language, per unit, 1990-2010 publication dates only
 - Counts of all monographs published 1990-2010 by unit library and language, sorted from most common to least common language. Each unit library on a separate worksheet tab. Music scores on a separate tab, with all locations on a single tab. Includes acquisitions through mid-April 2010. Excludes cons and cts locs. Maps grouped with Olin. Sublocations included with main location, except for annex locations, which are shown separately on the same tab as the associated main location. Excludes lost/missing/withdrawn items.
 - Includes a tab showing language breakdown of annex locations only, separately
 - Includes a tab showing consolidated annex and non-annex holdings, broken down by language, including percentages
 - Includes a tab showing percentage of total holdings in each language that are held in the annex
 - Includes a tab summarizing total holdings by unit (main locations and associated annex locations)

- D) Holdings by LC top class (first letter of LC class), per unit, 1990-2010 publication dates only
- Counts of all monographs published 1990-2010 by unit library and LC top class. Each unit library on a separate worksheet tab. Music scores on a separate tab, with all locations on a single tab. Includes acquisitions through mid-April 2010. Excludes cons and cts locs. Maps grouped with Olin. Sublocations included with main location, except for annex locations, which are shown separately on the same tab as the associated main location. Excludes lost/missing/withdrawn items.
 - Includes a tab showing LC top class breakdown, annex locations only, separately, but in one long list
 - Includes a tab showing consolidated annex and non-annex holdings, broken down by top level LC class, including percentages
 - Includes a tab showing percentage of total holdings in each LC top class that are held in the annex
 - Includes a tab summarizing total holdings by unit (main locations and associated annex locations)

II) Circulation frequency, items published 1990-2010, per unit

A) by LC subclass and publication date

- Historical circulation stats for monographs (books only) published 1990-2010, by unit library, publication year, and LC subclass, as of mid-April 2010. Each unit on a separate tab. Excludes cons, cts, and all non-circulating locations Sublocations included with main locations, except for annex locations, which are included with their associated main location. Excludes lost/missing/withdrawn items
- Includes separate tab for music scores, all locations
- Includes tab with all locations on a single tab, in one long list
- Each worksheet each unique combination of unit, publication year, and LC subclass, includes total volumes, total circulated volumes, total circulation, average circulations for circulating volumes, average circulations for all volumes, and percentage of volumes that have circulated.
- Tab labeled "olin2circs" includes two bar graphs showing Olin monograph circulation trends for items published 1990-2010, by publication year. One includes Olin annex holdings, the other does not. The one excluding annex holdings shows an unusual bump in circulation for items published 1990-1992, but this is due to the removal of large number of low circulation items to the annex for those publication years.

B) by LC subclass and language

- Historical circulation stats for monographs (books only) published 1990-2010, by unit library, LC subclass, and language, as of mid-April 2010. Each unit on a separate tab. Excludes cons, cts, and all non-circulating locations Sublocations included with main locations, except for annex locations, which are included with their associated main location. Excludes lost/missing/withdrawn items
- Includes separate tab for music scores, all locations
- Includes tab with all locations on a single tab, in one long list
- Each worksheet, for each unique combination of unit, LC subclass, and language, includes total volumes, total circulated volumes, total circulation, average circulations for circulating volumes, average circulations for all volumes, and percentage of volumes that have circulated.
- Tab labeled "olin2circs" includes two bar graphs showing Olin monograph circulation trends for items published 1990-2010, by language of publication. One shows the percentage that circulated as of mid-April 2010 for the top twenty languages according to

holdings volume. The other shows the same data, but for the top twenty languages with the highest percentage of circulating volumes, regardless of holdings status, except that it excludes languages that had fewer than 100 total volumes in the 1990-2010 set.

C) by publication year only

- Historical circulation stats for monographs (books only) published 1990-2010, by unit library, and publication year, as of mid-April 2010. Each unit on a separate tab. Excludes cons, cts, and all non-circulating locations Sublocations included with main locations, except for annex locations, which are included with their associated main location. Excludes lost/missing/withdrawn items
- Includes separate tab for music scores, all locations
- Includes tab with all locations on a single tab, in one long list
- Each worksheet, for each unique combination of unit, LC subclass, and language, includes total volumes, total circulated volumes, total circulation, average circulations for circulating volumes, average circulations for all volumes, and percentage of volumes that have circulated.

III) Circulation frequency, all items, per unit

- This was intended to be a series of files containing data formatted as in II) above, but for the entire collection. However, data for pre-1990 publication dates could not be assembled in time for this report. It will be available at a later date.

IV) Circulation of monographs acquired in 2001, per unit

A) By year of circulation, total, and by patron group

- Circulation of monographs acquired in 2001 from Jan 1, 2001 until April 26, 2010, showing total circulation over that time, plus breakdowns by year of circulation, and by a modified patron group classification. One tab for each unit library. Excludes lost/missing items, and cons and cts locations. Sublocations consolidated under main locations, except for annex locations, which are listed with the associated main location. Maps locations changed to Olin. Excludes internal use circulation (special projects, digitization, new book shelf, etc.) Excludes non-circulating location items.
- Includes separate tab for music scores, all locations
- Tab labeled "olin2circs" includes a number of bar graphs showing circulation trends for Olin holdings acquired in 2001. One set of graphs shows the percentage that circulated during a particular year from 2001 until 2009 (one graph per year), broken out by LC top class (first letter of LC classification) as of mid-April 2010. There is also a graph showing average circulation percentage per LC top class across the entire period from 2001-2009. The other set of graphs shows the percentage that circulated for a particular year LC top class (one for each top class except for 'M', since Olin acquired only 2 'M's in 2001), broken out by year of circulation, as of mid-April 2010. There is also a graph showing average circulation percentage for all LC top classes (that is, for the entire set of titles acquired by Olin in 2001), again broken out by year of circulation.

V) Circulation of monographs published 1990-2010

- Historical circulation stats for monographs (books only) published 1990-2010, as of mid-April 2010. Each unit on a separate tab. Excludes cons, cts, and all non-circulating locations Sublocations included with main locations, except for annex locations, which are included with their associated main location. Excludes lost/missing/withdrawn items
- Each tab shows circulating volumes (the number of volumes in the category that circulated at least once since acquisition (assumed to be near date of publication), total

volumes (the total number of volumes in the category) and percentage circulation (circulating volumes divided by total volumes, expressed as a percentage, and shows the percentage of volumes in the category that circulated at least once since acquisition/publication)

- Includes a tab arranged by publication year and LC subclass, consolidated across all units.
- Includes a tab arranged by publication year and language of publication, consolidated across all units
- Includes a tab arranged by unit library, publication year, and LC top class (first letter of LC class). All units on one worksheet.
- Includes a tab arranged by unit library, publication year, and language of publication. All units on one worksheet.

VI) Snapshot reports (from April 19, 2010 snapshot)

A) Circulation by unit and LC subclass; circulation by unit and language arranged by patron status

- The circulation snapshot captures data about all items that are charged out as of the date of the snapshot. It combines bibliographic data from Voyager with patron data from Human Resources and other sources. Thus it allows closer tracking of the patron status and academic affiliation of users of the print collection, at least for items charged out to current, active members of the Cornell community. The patron group tracking possible with the circulation snapshot is more fine-grained than that recorded in Voyager's circulation transaction archive. For example, it permits a distinction to be made between graduate carrels and faculty studies, between traditional ILL and Borrow Direct, and between academic and non-academic staff
- Includes tab showing circulation (on 4/19/2010) by unit, LC subclass, and patron group for books, and a separate tab for scores
- Includes tab showing circulation (on 4/19/2010) by unit, language, and patron group for books, and a separate tab for scores

B) Circulation arranged by academic affiliation

1) By unit and LC subclass

- Circulation counts from the snapshot of 4/19/2010 (see above) with each unit library and LC class broken out by academic affiliation. Academic affiliation classification provided in the Human Resources system varies according to status. Therefore, circulation data correlated with academic affiliation is provided on different tabs, depending on patron status.
- Faculty by college or school affiliation
- Faculty by departmental affiliation
- Graduate students by field of study
- Graduate professional school students by college or school affiliation
- Undergraduate students by college or school affiliation
- Academic staff by departmental affiliation
- Non-academic staff by departmental affiliation
- Separate tab for music scores, all patron groups

2) By unit and language

- Same as above, except substitute language for LC subclass

VII) Unique item circulation for monographs acquired in 2001, per unit

- This differs from IV) above. IV tracks circulation year to year within a subset of volumes, but does not consider which volumes have circulated. This worksheet tracks

unique item circulation year to year. Each time a volume within a set circulates for the first time, it is added to the tally for that set. So the circulation totals for each year are cumulative, and include tallies for volumes that circulated in previous years. The line graph that can be seen by scrolling to the right shows the growth in unique volume circulation for each unit library collection for items acquired in 2001, over the years from 2001 to 2009. Units differ with regard to both the rate at which previously non-circulating titles continue to be checked out over time, and the point at which use of new volumes starts to level off.

- Beneath the line graph is a bar graph. This shows the number of different years in which the average volume in a unit library collection circulated, for 2001 acquisitions, over the period from 2001-2009. Thus, the maximum value for this measure would be nine, if every volume in a collection circulated at least once in every year between 2001 and 2009. The lowest possible value is one, for a collection where every circulating volume circulated once and only once over the nine year period. (Completely non-circulating volumes are not included in this analysis). The higher the value for a collection, the greater the tendency for repeat circulation of volumes over a period of years. Thus, this is a partial measure of the staying power, or popularity of volumes in a collection over time. The measure is an average, and does not distinguish between a collection that may have a core of titles that circulate very regularly from a collection that has a higher number of volumes that circulate less often.
- Excludes 2010 data (a partial year) to avoid creating an artificial rolloff. All sublocations are rolled up to main locations except annex. Cons and cts locations excluded, and maps location is converted to Olin. For book only (does not include music scores). Excludes lost/missing/withdrawn items. Excludes internal charges, including special projects, digitization, and circulation to new book shelves. Excludes charges for items in non-circulating locations.