

Purdue University
Purdue e-Pubs

Charleston Library Conference

Are E-Book Packages Overwhelming and Redefining Your Collection?

Rhonda R. Glazier
University of Colorado at Colorado Springs, rglazier@uccs.edu

Matthew J. Jabaily
University of Colorado, Colorado Springs, mjabaily@uccs.edu

Lynn E. Gates
University of Colorado, Colorado Springs, lgates@uccs.edu

Follow this and additional works at: <https://docs.lib.purdue.edu/charleston>

 Part of the [Collection Development and Management Commons](#)

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

Rhonda R. Glazier, Matthew J. Jabaily, and Lynn E. Gates, "Are E-Book Packages Overwhelming and Redefining Your Collection?" (2017). *Proceedings of the Charleston Library Conference*.
<http://dx.doi.org/10.5703/1288284316684>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Are E-Book Packages Overwhelming and Redefining Your Collection?

Rhonda R. Glazier, University of Colorado Colorado Springs

Matthew J. Jabaily, University of Colorado Colorado Springs

Lynn E. Gates, University of Colorado Colorado Springs

Abstract

Most academic libraries offer numerous e-books alongside their print titles. Traditionally, print materials have been chosen by subject liaisons with input from departmental faculty, whereas e-books have been acquired en masse through large collection purchases, subscriptions, or PDA/DDA programs that include large numbers of discovery records. At Kraemer Family Library, the print budget is divided into subject areas using a formula that includes the number of students in a discipline, level of program (bachelor, master, or doctorate), number of faculty by discipline, and average cost of materials in a subject area. This budget breakdown is an attempt to balance the library's collection so that it reflects the focus and emphasis of the curriculum being taught on campus. Beginning in 2012, the Kraemer Family Library at the University of Colorado Colorado Springs began participating in two PDA/DDA e-book programs. The library also began purchasing e-book packages that were either publisher or subject based. During this same time, the library continued to use a formula to allocate the library's print budget. Because e-books were not purchased according to any allocation, and the library was beginning a process of weeding the print collection, an analysis of the effect of e-books on the overall collection was undertaken. The purpose of this analysis was to determine what metrics should be used to determine the impact of e-books on the overall collection and to analyze that data for overall impact on the collection.

Background

The University of Colorado Colorado Springs (UCCS) is one of four campuses in the University of Colorado System. In the fall of 2017, UCCS had 12,400+ students enrolled. There are currently over 790 faculty members teaching in 46 baccalaureate degrees, 22 master's, and 5 doctoral programs. The Kraemer Family Library (KFL) at UCCS began participating in two DDA/PDA programs in fiscal year 2012. One DDA program was through the Colorado Alliance of Research Libraries and is publisher based. The second program is profile based and is administered by the CU Boulder campus. Around the same time, the library began purchasing e-book packages that were either publisher- or subject-based collections. Bibliographic records for these titles are added to the catalog. In 2016, the library added ProQuest's discovery layer, Summon, to the catalog and began teaching this new search tool in library instruction. The search bar is also prominent on the homepage of the library's website and is the primary mechanism for students and faculty to find materials at the library. During this time the library continued to allocate the print budget based on a formula that included: number of degrees offered by discipline; number of faculty by discipline; credit hours generated and weighted by level; and average cost of books by subject. This formula is an attempt to

balance the collection so that it reflects the priorities and curriculum on campus. There was no attempt during this same time to shape the e-book collection in the same way.

After many discussions about the increased emphasis on e-books in the collection, we decided that a method to effectively analyze the overall collection needed to be developed. This analysis needed to include both print and e-books. The following questions formed the basis for our analysis:

- What is the impact of the e-book collections on the scope, depth, and valance of the overall collection?
- Does coverage and use differ between e-book and print collections?
- Are the e-books impacting the effectiveness of the print collection?
- What data should be used to analyze a collection consisting of both e-books and print books?

After determining the questions that we wanted to answer, we then needed to determine the best way to find the answers. While there are many ways to

analyze a collection, we finally decided to look at the following areas:

- Budget
- Number of print vs. e-book titles added over the past five years
- Subject distribution by LC call number and fund code
- Usage of print titles and e-books added over the past five years
- Effectiveness of the catalog to retrieve print and e-book titles

Budget

The breakdown of the Kraemer Family Library's budget shows that it is very similar to the budgets of other libraries. In FY17, KFL spent 72.63% of the materials budget on electronic resources and 7.66% on print serials, leaving 19.71% to spend on single title print monographs. This 19.71% is divided between subject areas using a budget formula that includes:

- Percentage of programs by department
- Percentage of credit hours generated weighted by level (freshman, sophomore, junior, senior, graduate)
- Percentage of faculty by department

These elements are averaged to determine demand and added to average percentage of the cost of a book by subject area. This formula is intended to balance the collection according to the campus curriculum and research emphasis. This "print" budget has broadened over the years to include DVDs and other single title purchases. Beginning in FY17, the library started purchasing single title e-books from the monographic print budget. However, at this time it is still a very small number. When adding e-books to the library's collection, there was no allocation by subject. The cost of participating in two DDA programs and any subject or publisher e-book packages were taken off the top of the budget, with no attempt to balance the spending by subject. Opportunities to purchase e-book packages that were deemed useful and "a good buy" were added with little attempt to balance the content with other e-book packages. This method of adding e-books as the opportunity arose has continued over the five years that the library has been adding e-book titles to the collection.

Print vs. E-Book Titles

Although it was known that a large number of e-books from packages and discovery records had been added to our system, the exact number had not been calculated. To study this, we created review files of all print and e-book bibliographic records using Sierra, the library's integrated library system. For the purpose of this study, the list was limited to titles obtained in the last five years. For the print list, only titles that were purchased (no gift books) and in the main collection were counted. For e-books, all titles that were entered into the catalog during the past five years were counted, including purchased e-books and DDA/PDA discovery records.

The results showed that about 25,000 print books and nearly 100,000 e-books were added to the collection in the last five years. The ratio of e-books to books was consistent over the five-year period at about 4:1. This is in spite of spending more money on print titles than e-books. Although our older collections are still heavily weighted toward print, this imbalance in newly added titles has shifted the overall balance of our collection, and we expect the trend to continue.

Call Number and Subject Distribution

The next step was to determine the amount of impact this shift had on the subject balance of the collection. Using the previously generated lists, call numbers were exported for the print and e-books separately. Call numbers were assigned to e-books that were missing them when possible. The call numbers were then truncated to the first two letters, sorted, and counted in Microsoft Excel, giving a list of the number of titles for each call number area for both print and electronic books. The number of print and e-book titles in each call number area was then divided by the total number of print or e-book titles to determine the percentage of the respective collections in each call number area.

The results showed a wide discrepancy in many call number areas. Print books were relatively evenly distributed across call number areas compared to e-books, which were more concentrated in fewer call number areas. For example, QA (mathematics) titles were the most numerous for both print and e-books, but for e-books QA titles were more than 10% of the collection compared to less than 6% of the print collection. Books with call numbers starting in TA and TK (general and electrical engineering)

were also heavily weighted toward e-books. On the other hand, some call number areas, especially PN and PS (general and American literature), were weighted heavily toward print.

While initial analysis provided a wide look at the collection, a second analysis looked at four specific academic areas in depth. The four areas selected were business/economics, education, electrical engineering, and psychology. They were chosen because they each represent an important UCCS program, had sufficient collection size, and were diverse in terms of print/electronic discrepancy.

One problem with analyzing collections according to coverage in academic area is that tools for doing so are limited. While call numbers classify titles according to subject, they can be imprecise because each title is limited to a single call number but may cover many subjects. Academic areas are also diverse in their needs for materials from a variety of subjects. For print titles, UCCS librarians select individual titles and designate the academic area for which they are purchased by designating a fund code. This allows the library to track which titles should be associated with which academic areas for analysis purposes. Unfortunately, e-books are not selected individually or purchased using fund codes associated with academic areas, so they cannot be analyzed in this way.

Because of these limitations, the analysis of the collection by academic discipline was done using call numbers. An additional analysis was done to see to what extent the purchases of print books by fund aligned with associated call number ranges. Comparing lists of print titles purchased by the four academic discipline funds to titles purchased in the associated call number areas, a varying degree of overlap was found. The percentage of titles purchased with academic area funds that were within the associated call number ranges were as follows: Business/Economics 81%, Education 75%, Electrical Engineering 53%, and Psychology 78%. Conversely, the percentage of titles in academic discipline call number ranges that were purchased with the associated funds were as follows: Business/Economics 75%, Education 88%, Electrical Engineering 59%, and Psychology 77%. These results suggest that measuring academic discipline by associated call number range is reasonable but does not necessarily tell the whole story.

Returning to the question of balance, the results showed that print and e-books represented

reasonably similar proportions of their respective collections in some academic areas but not others. Business and economics titles were a large part of both print and e-book collections, but the e-book titles were a slightly larger proportion of the electronic collection. Representation in psychology was relatively even. In education and electrical engineering, however, there was a significant imbalance. Education print titles have a much higher proportion of the print collection than of the electronic collection. Electrical engineering titles, on the other hand, accounted for a small portion of the print collection but a much larger portion of the electronic collection.

A final analysis focused on usage of print and e-books. For the print collection, circulation data was pulled from the library's integrated library system. The number of checkouts were counted and tallied for the call number ranges used in the study. They were then divided by the total number of checkouts to determine the percentage of all checkouts from each subject area. The results were compared to the percentage of titles that were from that academic area. This allowed the determination of whether a book in that area was more or less likely to circulate than the average book.

A similar analysis was done for e-books, but limitations in reporting made things difficult. As a result, only e-books from one provider, Springer, were analyzed. Springer was chosen because it was the library's largest e-book collection and had consistent usage reports for the past five years. The usage reports were matched with the exported data from the integrated library system, allowing the determination of chapter downloads in each call number range. Again, a percentage was determined by dividing by the total number of downloads. As with the print results, the percentage of titles and downloads were compared to see whether e-books from the academic areas were more or less likely than the average e-book to be accessed.

In business/economics, print titles were more likely to circulate than the average book but e-books were less likely than the average book to be downloaded. In education, both print and e-books underperformed the average book. In electrical engineering, print books were checked out in proportion to their representation in the collection, but e-books dramatically overperformed in terms of downloads. Psychology books and e-books were both very popular, outperforming the average book in terms of checkouts and downloads.

Effectiveness of Catalog

In order to test the effectiveness of our catalogs, we wanted to compare our two catalog products that our patrons use to find materials. Our classic catalog is III's Sierra WebPAC and until late 2016, it was our only catalog. It is a traditional OPAC and there is no delay between when records are loaded, deleted, or changed and when they can be viewed online. However, not every collection we subscribe to is loaded into the catalog, and while the catalog is available from our homepage, there are extra clicks required to access it. The other way to access our collection is through OneSearch, our branded instance of ProQuest's Summon. OneSearch is the default search on our homepage and is taught in our information literacy and instruction sessions. However, unlike the WebPAC, information on books and e-books is not live; new, updated, and deleted records are pulled from the catalog weekly and sent to ProQuest.

We knew that for testing we wanted to use actual searches. While we would be unable to determine if a search was successful according to the original searcher, we would be searching the WebPAC and OneSearch the same way our patrons search, not the way a librarian would search. Google Analytics is used by our Web Services and Emerging Technologies Division to collect various pieces of information on both of our catalogs. They were able to provide the actual search performed in the catalog and were able to export 5,000 searches at a time to Excel. There were two things we needed to be aware of about our catalogs as we started a very small-scale study. One, many of our e-books use vendor records; these records tend to be brief with few, if any, subject headings or notes, such as a table of contents or summary. Two, in addition to the records that we load into OneSearch, many of the collections have been turned on in SerialsSolution. Thus, there are extra search points in the SerialsSolutions metadata to search against OneSearch. This extra metadata is not available to the WebPAC, which can only search what is in the record.

In performing the test searches, we used patron searches from both the WebPAC and OneSearch and each search was replicated in both catalogs. The top ten results were reviewed for type of material (book vs. e-book) and a title list was created. In the test searches, we found that e-books outnumbered books in the top ten search results in the catalog, except for the search that had a date limit of

1980–1989. In Summon, the e-book and book results were fairly balanced, but frequently journal articles would overwhelm both. One unexpected result was that despite doing the same search in both the classic catalog and Summon, there were very few titles that were in both search results list. The results from this small-scale search have shown us that we need to do a larger study.

Conclusions

Our study of our e-book collections was informative and surprised us in a few ways. We can say with certainty that the overall balance of our collections has shifted significantly. While we expected to see a growth in e-books, the 4:1 ratio of e-books to books in new titles added was larger than we expected.

Our call number–based analyses also showed that e-books were balanced differently than the books in our print collection. In some areas, like electrical engineering, this imbalance appears to be desirable. Our electronic holdings supplemented a relatively small, expensive collection with high-demand, low-cost titles. In other areas, like business and economics, it appears we may have added many low-demand titles to an already healthy print collection. We also discovered that some areas in our collection, especially in the humanities, have very few e-books.

Due to limitations of our analysis (reliance on call numbers, limited usage data, etc.) we must be careful not to overinterpret our findings. Nonetheless, our findings have suggested to us some possible courses of action. We may want to bolster areas where we have already seen success, continuing to develop our e-book collections in science and technology. On the other hand, we should also consider intentionally expanding our e-book collection into the humanities. We may also look into the budgeting process, revisiting our formula for allocating print funds in light of our e-book collections.

With regard to the analyses of our catalog and discovery layer, it did not appear that e-books were completely overwhelming the print collection, though the balance was more heavily weighted toward e-books in the WebPAC catalog. It was also apparent that the two platforms employed very different search algorithms, as their results had very few overlapping titles in the top results. These preliminary results suggest to us that a more comprehensive study needs to be done on this topic.