

12-4-2018

Documenting the Stewardship of Libraries: The Eastern Academic Scholars' Trust Validation Sample Studies

Sara Amato

Eastern Academic Scholar's Trust, samato@blc.org

Susan Stearns

Boston Library Consortium, sstearns@blc.org

Follow this and additional works at: <https://digitalcommons.du.edu/collaborativelibrarianship>

Recommended Citation

Amato, Sara and Stearns, Susan (2018) "Documenting the Stewardship of Libraries: The Eastern Academic Scholars' Trust Validation Sample Studies," *Collaborative Librarianship*: Vol. 10 : Iss. 3 , Article 4.

Available at: <https://digitalcommons.du.edu/collaborativelibrarianship/vol10/iss3/4>

This From the Field is brought to you for free and open access by Digital Commons @ DU. It has been accepted for inclusion in Collaborative Librarianship by an authorized editor of Digital Commons @ DU. For more information, please contact jennifer.cox@du.edu, dig-commons@du.edu.

From the Field

Documenting the Stewardship of Libraries: The Eastern Academic
Scholars' Trust Validation Sample Studies

Sara Amato (samato@blc.org)
Data Librarian, Eastern Academic Scholars' Trust

Susan Stearns (sstearns@blc.org)
Executive Director, Eastern Academic Scholars' Trust

Abstract

The Eastern Academic Scholars' Trust ([EAST](#)), a regional shared print collaboration of sixty academic and research libraries, conducted validation studies of the collective monograph collection in 2016 and 2017. Methodology, results and limitations of the studies are presented along with thoughts on further research areas.

Keywords: validation, shared print

Background on EAST and Validation Studies

The Eastern Academic Scholars' Trust (EAST) is a regional collaboration of sixty academic and research libraries to define and manage retention agreements for scholarly publications in support of teaching, learning, scholarship, and research. EAST's membership, while primarily in the northeast of the United States, includes libraries as far south as Florida and as far west as Tennessee, with collection sizes ranging from just over 100,000 holdings to collections of over four million. All materials retained for EAST are held in place by the partner libraries, though some belong to smaller cooperatives which have off-site storage facilities.

Formed in 2015 with funding from the Andrew W. Mellon and Davis Educational Foundations,¹ validation of shared print monograph retention commitments has proven problematic due

EAST addresses directly the growing need for academic libraries to ensure that monographs and journals of scholarly value are not inadvertently discarded as libraries undertake weeding and deselection programs to free up space for other services. EAST Retention Partners commit to retain titles of scholarly importance in their local collections for a minimum of fifteen years and make those titles available to faculty, scholars, and students at other EAST libraries. EAST is focused on sustained cooperation and trust across the partner libraries, and the validation studies described below have played a major role in instilling that trust.

While many serial and journal shared print programs engage in various levels of validation, e.g., the Western Regional Storage Trust (WEST)



to the large number of physical items that would require validation and limited resources to undertake the work. A notable exception to this is the Central Iowa Collaborative Collections Initiative (CI-CCI), which validated all 144,000 of their retention commitments in 2014. Their work, along with a review of other shared print monograph validation programs, was published in 2016². While full scale validation is laudable, it was not feasible with EAST's over nine million retention commitments.

In order to better understand the reliability of the EAST collective monograph collection and to help establish trust across the partner libraries, EAST undertook a project to design, test, and analyze a sample-based validation study. The purpose of this study was to determine the statistical likelihood that a retained title actually existed on the shelves of retention partners. The study was conducted with the initial forty Cohort One libraries during the spring of 2016 and was repeated with a second cohort of twelve libraries in the fall of 2017. Libraries participating in the study were compensated from funding provided by the Andrew W. Mellon Foundation. The validation sample studies, as well as the other work of EAST, was coordinated through EAST's administrative and fiscal host, the Boston Library Consortium (BLC).

Goals & Methodology

The goal of the Validation Sample Study was to determine the likelihood that an individual volume would be on the shelves at its owning EAST library and be in usable condition. The dual constraints of time and money meant that the validation study would have to be a statistical sampling of the collections rather than a full validation of retained materials. EAST engaged Professor Grant Ritter, Ph.D., of Brandeis University, to be the statistician for this project. In consultation with Professor Ritter the study was designed to predict the likelihood that an item

which had been assigned a retention commitment in the local catalog was actually on the shelf. The study was designed to be accurate within a 1% margin of error. The study also attempted to assess the condition of the materials to determine if they were in good enough condition to circulate. Professor Ritter developed a methodological approach to the study and provided statistical analyses of the condition of monographs and identifying factors which might predict a higher risk for being missing or in poor condition.

In addition to the statistician and the EAST leadership team, a ten person Validation Working Group was formed to advise on implementation and to create and test training documentation. The working group consisted of staff from member libraries who work in collections and stacks management and had at least one representative for each of the major Integrated Library Systems in use in EAST. The Validation Working Group also had representatives from small, medium, and large collections.

Developing the Sample to be Validated

A sample size was determined by balancing the desire for 1% accuracy with the need to have a sample size small enough for the libraries to complete the physical validation in a relatively tight timeframe. EAST also wished to provide the libraries with adequate compensation for the work they undertook.

From a statistical perspective, the sample size to assure 1% accuracy does not depend on the number of monographs in the library but rather on the underlying rate of missingness at the owning library. (Note: this was a concept new to many involved in the project.) After discussion with the Working Group it was estimated that libraries were unlikely to have a missingness rate higher than 10%. Based on this assumption that the missing rate at a given library would be 10%

or less, and the agreement that we wanted to ensure that the missing rate was accurate within 1%, a sample size of six thousand titles per library was determined to provide statistical validity. Should an EAST library have a missing rate higher than 10% upon completion of the sampling, the accuracy of their number would be less statistically valid. However, as noted below, that did not occur.

The Working Group recommended a six thousand title sample size to the EAST's governing body, the Executive Committee, and it was approved.

The Validation Working Group also discussed how to record and count the status of materials. A book could either be present on the shelf or not. For materials not on the shelf, the library's online catalog could have a record of its status, e.g., it might be in circulation, in repair, checked out, long overdue, or already known to be missing. Given that the intent of the study was to determine if books in the collection analysis were actually available to circulate, the Working Group determined the following statuses would be used to record and count the items:

- Items that were present on the shelves OR identified as in circulation per the library's online system were deemed "Accounted For"
- Items whose status was lost, missing, billed, or unknown were deemed 'Unaccounted For' (these became affectionately known as the LMBO items).

Assessing Condition of the Items

In addition to status, the Validation Working Group also considered how best to assess the condition of materials. Given that the majority of the actual validation work was to be done by student workers in the stacks, this assessment needed to be both quick and consistent. Originally, a five point Likert scale was considered,

but was quickly determined to be both too difficult to rate quickly and consistently, and did not provide additional actionable data. The Working Group proposed a three point scale of Poor, Good, and Excellent and created a detailed matrix with defining characteristics of each of the three categories. One member of the Working Group, Anthony Fonseca of Elms College, produced a training video explaining criteria with examples for each of the three ratings.³

Once the sample size, status categories, and condition assessment were settled, the group considered how best to develop the sample of six thousand items for each of the EAST Cohort One libraries. The sample lists were generated by OCLC's [Sustainable Collection Services®](#) (SCS) who were simultaneously working on the collection analysis for the forty EAST Cohort One libraries. Libraries provided SCS with lists of locations that should be excluded from the validation, e.g., high density storage facilities or other closed stack areas that were already under tight inventory control. SCS then queried their database for every nth title, where n equaled the number of items in the library divided by six thousand. Since SCS uses a PostgreSQL database, the query results were returned in random order, thereby ensuring a random sample for each library. SCS also sorted the returned query lists into location and call number order and created spreadsheets which contained both the bibliographic data and the item level enumeration and barcode. This was intended to best ensure that those doing the validation would have easy access to the items in call number order for each location. As described below, including the item barcode both simplified the data collection and best ensured that the correct item was being examined as demonstrated by the student workers as shown in Figure 1.

The Data Collection Tool

There was some concern that simply distributing the spreadsheets with instructions on how

to code materials could result in a variety of data coming back and that it would be better to constrain data collection more tightly. There was also a desire to be able to verify that the correct book was actually located.

To facilitate and simplify the data collection, assure standardized language, and verify that the correct item was being checked (using the barcode provided), the EAST Data Librarian, Sara Amato, created a front end to Google Sheets using Google App Script.⁴ This validation tool was tested by the members of the Working Group with various browsers and devices and was eventually used by all libraries in the study. The code for the Validation Tool is available on GitHub.⁵ In most cases, students workers per-

formed the data collection, though some libraries used staff, and in one case, graduate library school students were employed.

Using the tool, the worker downloaded a specified number of items to validate. The items were presented in call number order. Upon locating an item, the worker scanned the barcode to indicate its presence on the shelf and then selected the condition option most appropriate to the item. Once the list of items was validated, it was uploaded and the validation dashboard, shown in Figure 3, was updated. This allowed the EAST Project Team to track each library's progress against the six thousand items required.

Figure 1. Students doing data collection in the stacks.



Execution and Challenges

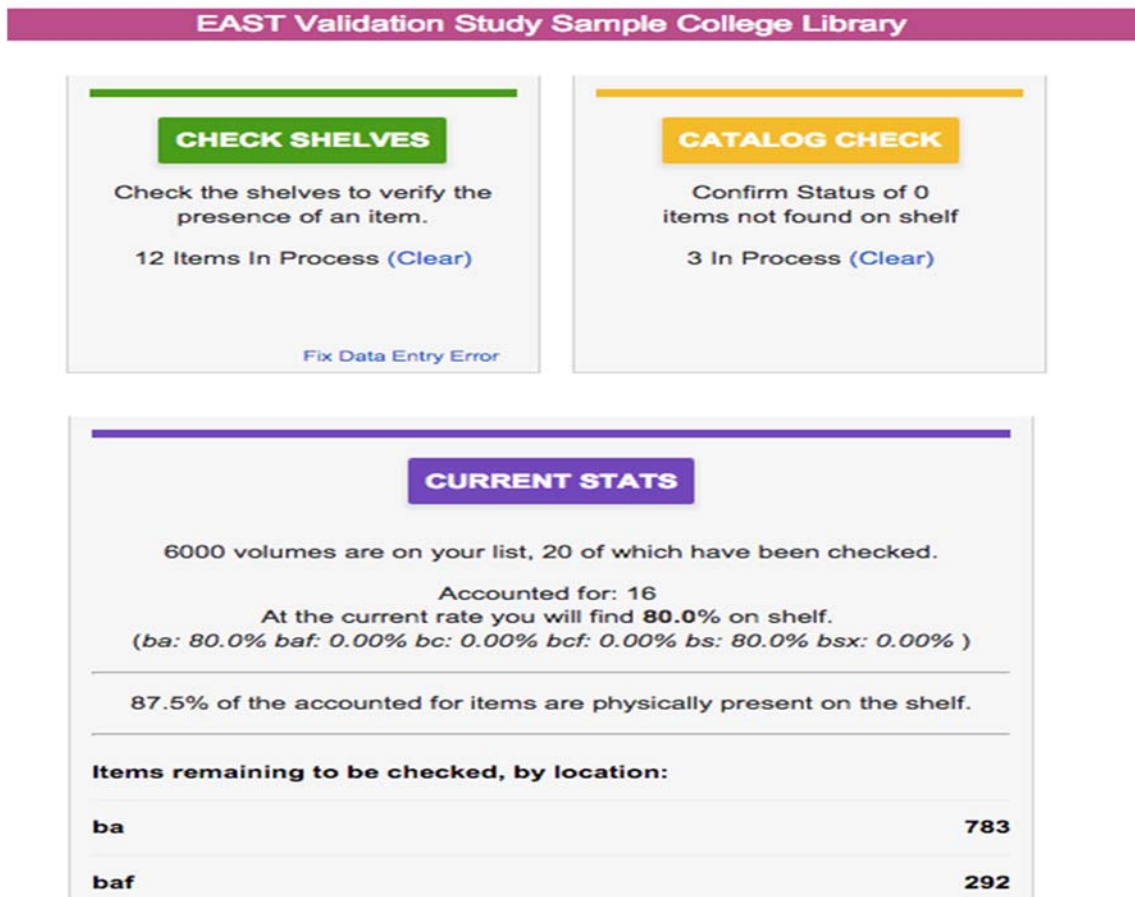
Preparation for the Data Collection

The two EAST Validation Sample Studies were conducted over a twelve week period in the spring of 2016 with the forty Cohort One libraries and in the fall of 2017 with the twelve Cohort Two libraries. Each library was tasked with validating the presence and condition of 6,000 titles randomly selected from those included in the collection analysis completed with SCS. The member libraries provided staffing to conduct the study and were reimbursed for labor and ad-

ministrative costs from grant funding. Each library was required to attend an initial kick-off/training webinar, after which they were provided with their validation list.

Before commencing the on-shelf validation, libraries compared their validation lists to the local ILS in order to record which items were currently checked out or known to be missing. This data was reported back to EAST and merged into the spreadsheet used by the validation tool. The library was then given a link to an instance of the validation tool which interacted with the Google Sheet that contained their library's data and could begin the data collection in the stacks.

Figure 2. Opening screen of validation tool.



Collecting the Data in the Stacks

Using the validation tool, workers could decide how many titles to validate in a session. They were then presented with the title and call number of the item to be validated, and asked to wand the barcode of the item. This provided an extra level of confidence that the item being examined was the correct item from the list. The workers were provided with immediate feedback if the barcode did not match and had the option to override the error if they were certain

they had the correct book. Wi-fi was not needed while doing the work in the stacks but was required to upload the results at the end of each session.

Results of each session were immediately recorded in the Google Sheet, which in turn updated statistics on a local library dashboard and an EAST project team dashboard where the progress of the study could be monitored in real time.

Figure 3. Dashboard at end of study with all 6000 titles checked and final scores. The “Animal” designation provided anonymity for the participating libraries.

| EAST Validation Sample Study 1 Dashboard | | | | |
|--|-----------|-----------------|---------------------|--------------------------|
| Anonymization | | | | |
| Animal | # Checked | % Accounted For | % Average Condition | % Average in Circulation |
| panda | 6000 | 96.8% | 2.8% | 0.8% |
| monkey | 6000 | 95.3% | 1.7% | 3.7% |
| lemur | 6000 | 91.6% | 2.1% | 5.3% |
| fox | 6000 | 94.9% | 2.2% | 2.8% |
| sheep | 6000 | 97.0% | 2.2% | 2.7% |
| wombat | 6000 | 96.7% | 2.3% | 3.7% |
| bat | 6000 | 94.3% | 2.4% | 0.9% |
| chupacabra | 6000 | 97.3% | 2.2% | 2.8% |
| grizzly | 6000 | 99.0% | 2.7% | 3.4% |

Overcoming Data Collection Challenges

While most of the data collection progressed smoothly, one library did experience some trouble with the tool which, after troubleshooting, was determined to be due to the laptop disconnecting and reconnecting to different wi-fi

hotspots in the library as the students worked in the stacks. Once the laptop was configured to only use the main library wi-fi the problems resolved.

Another challenge was that some of the titles listed in the spreadsheets were determined to

have been included in error, most often representing titles from non-circulating special collection locations. This was usually caught early in the process and the validation lists were rerun excluding those materials.

Reports from the participants in both cohorts indicated an average of 46 books could be checked per hour with a minimum of 29 and a maximum of 102. This is slower than the books per hour rate reported by CI-CCI of 132 per hour, perhaps due to the added condition validation criteria.⁶ CI-CCI asked libraries to simply assess whether or not a book was in good enough condition to circulate and allowed libraries to use their own criteria in making this assessment. In retrospect, this may have been a better option for condition assessment, as in the end EAST did not make any actionable distinctions between materials in good or excellent condition, but did use the data, as described below, to facilitate further statistical analysis.

The average administrative time needed was 25 hours which included running reports against the ILS, training, and supervising student workers. While these numbers are only estimates, they would tell us that replicating the study with a sample size of 6,000 would require around 130 hours of time validating in the stacks and 25 hours of administrative overhead for a total of approximately 155 hours.

Results

Cohort One Results - Missingness

All of the Cohort One libraries completed data collection by the deadline of April 22, 2016 and the raw data was provided to Professor Ritter

for analysis, the summary of which is in Appendix A. (Note that Appendix A contains the description results from Cohort One, Cohort Two, and a combined analysis which included both Cohort One and Cohort Two data.) Overall, 97% of monographs in the sample set were accounted for (mean), with a median of 97.2%. A summary of the distribution is in Table 1. All results were accurate within 1%. An average of 2.3% of titles were in circulation at the time of the study and 87% were validated by using their barcode.

Based on the statistical analysis completed by Dr. Ritter, the factors correlated with missingness (i.e., not being present on shelf or accounted for in circulation) included age, frequency of circulation, and certain subject areas such as Mathematics (QA) and U.S. Law (KF), though none of these were deemed significant enough to warrant action. The only factor which was statistically significant in its correlation with missingness was the owning library. Libraries with lower validation scores were often aware of internal problems that lead to higher rates of missingness or bibliographic error, such as one library that reported an inaccurate inventory procedure a decade earlier.

Cohort One Results - Condition

Of the titles in the Cohort One sample set, approximately 90% were deemed to be in average or excellent condition. Approximately 10% of titles reviewed were marked as being in poor condition. The mean score for condition was 2.2 (1=poor, 2=acceptable, 3=excellent).

Table 1. Summary statistics and distribution on missing monographs at the forty Cohort One monograph retention partners.

| | mean | std | Min | 5 th pctl | 25 th pctl | 50 th pctl | 75 th pctl | 95 th pctl | max |
|-------------|------|------|-----|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------|
| Rate | 3.0% | 2.1% | .3% | .6% | 1.3% | 2.8% | 4.2% | 7.4% | 9.7% |

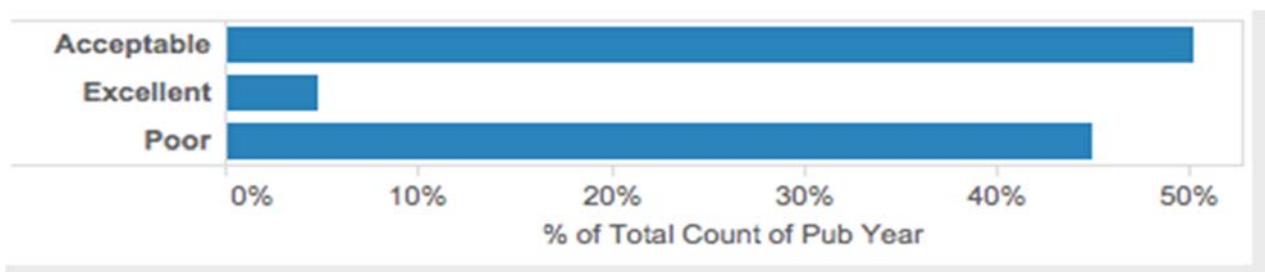
Table 2. Summary statistics and distribution on condition of monographs at the forty Cohort One monograph retention partners.

| | mean | std | Min | 5 th pctl | 25 th pctl | 50 th pctl | 75 th pctl | 95 th pctl | max |
|-------------------|-------|-------|-------|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-------|
| Poor | 10.7% | 9.8% | 1.1% | 1.3% | 3.0% | 8.2% | 15.2% | 29.3% | 44.2% |
| Acceptable | 55.7% | 17.3% | 34.8% | 34.8% | 45.5% | 54.1% | 65.9% | 88.1% | 91.1% |
| Excellent | 36.6% | 17.7% | 6.8% | 7.8% | 20.7% | 33.6% | 43.3% | 70.5% | 78.8% |

While some subject areas were determined to be in slightly worse condition, for example monographs on Paintings and African History (call numbers ND and DT respectively), the more significant factors were use, age, and to some extent location. Each increment of twenty additional checkouts increased the likelihood of being in poor condition by 5.1% (for example, a

10% likelihood would become a 10.51% likelihood). Each increment of ten years in the age of the item increased the likelihood of being in poor condition by 3%. Since some items have recorded uses in the hundreds and some items are greater than 50 years old, both of these factors could predict much higher likelihoods of poor condition.

Table 3. Condition of titles published before 1900. Older titles are in worse condition.



There was also significant variability in condition ratings at different holding libraries. While the Validation Working Group did produce training tools for assessing condition, there was some concern over the consistency of the ratings as they were, in the end, subjective.

Further Analysis of the Validation Data Set

The Validation Sample Study with Cohort One libraries took place concurrently with the collection analysis being done with SCS. Unfortunately, this meant that the results of the validation study were not yet available to inform the collection analysis model being used to determine which titles libraries would agree to retain for 15 years (referred to as the retention model).⁷ Fortunately, the final retention model did include keeping a higher number (up to five) of highly used titles, even without knowing that these were at a higher risk of being in poor condition. In addition, of the 240,000 items sampled in the validation study, 92,575 subsequently received retention commitments, providing a large enough sample of the 'collective collection' to do statistically valid predictive modeling across the full set of EAST retained titles following the completion of the EAST collection analysis and retention allocations.

Using data from the validation study and data on the full holdings of the forty EAST Cohort One libraries, Professor Ritter identified 77,925 titles (.01% of the collective collection at that time) as having a greater than 7.5% chance of being missing or a greater than 50% chance of being in poor condition. These were titles with only one copy being retained by EAST which had unallocated surplus copies at other EAST libraries. These titles were then provided back to the member libraries holding surplus copies with the highest validation scores as potential additional retention candidates. The majority of these additional copies of at risk titles were accepted as additional retention commitments at Cohort One libraries.

Some 5,000 of the 77,000+ titles were not accepted by the owning library as additional retentions for various reasons (e.g., not on shelf or the owning library was not willing to take on additional retentions), and approximately 9,000 of the titles determined to be at risk had no surplus copies in EAST Cohort One. These were eventually passed on as potential retention candidates in Cohort Two, where most were able to be retained.

This is, to our knowledge, the first time that validation sampling has been used to inform retention modeling for shared print. It is hoped that this type of data can help inform future shared print retention models.

Cohort Two Results

With the formation of a second cohort of libraries joining EAST, the opportunity arose to replicate the validation study. Cohort Two consisted of twelve libraries, with a much greater variability in size, ranging from collections under 200,000 to over 4 million. One library, Union College, had participated in EAST Cohort One as a contributing member, extending their ConnectNY retentions to EAST. As such, they did not participate in the collection analysis, but did complete the validation study using only their retained titles. In Cohort Two, they participated fully in the collection analysis and conducted the validation study again, this time against their in-scope titles in the collection analysis. This provided the opportunity to compare the results of their retained titles against the full collection.

Cohort Two libraries conducted the study in October through December 2017. Overall results showed 97.8% of monographs in the sample set were accounted for (mean: 97.8%, median: 98.15%, high of 99.8% and low of 94%). All results were accurate within 1%. 2.4% of titles were in circulation at the time of the study, and 95% were validated by using their barcode. The

similarity of these results to Cohort One results were reassuring to the Collection Analysis Working Group that materials are, in general, 97% likely to be available. Results for Union College were also very similar to the results of their first validation study, which gives some reassurance that validations over the entire collection vs. just retention commitments are analogous.

Factors correlating with missingness again included age, frequency of circulation, and certain subject areas though this time Religion (BL) was included in addition to US Law (KF).

Not surprisingly, age and circulation continued to be the most significant risk factors for an item to be in poor condition. Knowing this to be the case, Cohort Two chose a retention model that increased the number of retention copies of older materials. More specifically, they chose to retain up to three copies of titles published before 1900 when available. This, too, is an excellent example of how the results of the validation sample study were integrated into subsequent decisions about retention.

Table 4. Summary statistics and distribution in rates of missingness among 12 EAST Cohort Two monograph retention partners.

| | mean | std | 5 th pctl | 25 th pctl | 50 th pctl | 75 th pctl | 95 th pctl |
|-------------|-------|-------|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Rate | 2.18% | 1.55% | .25% | .86% | 1.85% | 2.89% | 5.42% |

Table 5: Distribution of monographs in poor condition among additional 12 EAST Cohort Two monograph retention partners.

| | mean | std | 5 th pctl | 25 th pctl | 50 th pctl | 75 th pctl | 95 th pctl |
|-------------|------|------|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Poor | 7.1% | 5.9% | .68% | 2.20% | 5.55% | 11.2% | 18.8% |

A comparison of the results of the two cohorts is shown in Table 6 below.

Table 6: Combined raw results from Cohorts One and Two.

| Cohort | # Checked | % Ac- counted For | Average Condi- tion 1=poor,2=aver- age,3=excellent | % Poor | % Validated by Barcode | % In Circula- tion |
|-----------------|----------------|----------------------|---|------------|---------------------------|-----------------------|
| One | 240,000 | 97.0% | 2.2 | 10 | 87.1 | 2.3 |
| Two | 72,000 | 97.8% | 2.4 | 7.1 | 94.9 | 2.4 |
| Combined | 312,000 | 97.2% | 2.3 | 9.3 | 89 | 2.3 |

Analysis of Combined Cohort One and Cohort Two Datasets

Perhaps most interesting was the opportunity to look at the combined datasets to see if any generalities could be drawn on factors influencing condition and rates of missingness. Again, Professor Ritter did the analysis for EAST and provided the following insights.

The data show mean estimated rates of missing equal to 3.0% for Cohort One libraries, 2.18% for Cohort Two libraries, and 2.79% for the combined group, giving confidence to an assumed 97% availability rate. Notably, all participating libraries had estimated rates of missing under 10.0%, indicating strong likelihood that all estimated rates of missing were accurate to within 1.0% and confirming our assumption that the libraries involved in EAST have, for the most part, been trusted curators of their local collections.

The only consistently significant predictors for an item being missing were the age of the monograph and having its subject matter classified as Religion ('BL') or US Law ('KF'). In particular, US Law monographs were 4.5% more likely to be missing and Philosophy and Religion monographs were 1.8% more likely to be missing. In addition, every ten year increase in the age of an item correlates with a 4% to 5% increase in its likelihood of being missing. Since some monographs are greater than fifty years old, this factor could predict a much higher likelihood (e.g., 20%-25% higher) compared with fairly new items. However, both data sets were consistent in noting the strongest characteristic increasing the likelihood of a monograph being missing was the library itself.

The results for condition rates have a much larger number of significant predictors. Based on Cohort Two results, age, circulation and number of US holdings were all associated with a higher likelihood of being in poor condition. In addition, Cohort Two libraries had materials

in poor condition in more subject areas than the Cohort One libraries. For example, monographs in the areas of Psychology, Asian History, Economic History, Family and Marriage, Welfare and Criminology, Theory and Practice of Education, Painting, and French and Spanish Literature were all more likely to be in poor condition. Almost all of these factors were also similarly significant when using the combined sample of Cohorts One and Two. Only two factors, number of US holdings and being classed with French or Spanish literature (call number 'PQ'), did not retain their significance with the combined sample. At the same time, only one factor, being classed with English Literature (call number 'PR') gained significance in the larger combined sample. For a more detailed analysis of these relationships see Appendix A.

One caveat to all of the analyses of monograph conditions is that the assessments at different libraries were done by different reviewers. Although reviewers were given training, it is likely they employed somewhat different standards for determining monograph condition.

Conclusions and Further Work

For the most part, the validation studies confirmed that the likelihood of a monograph being missing is low, reconfirming a 97% likelihood of material being available as was found with the Central Iowa Collaborative Collections Initiative (CI-CCI) study.⁸ There is, however, variation in the likelihood of being missing based on the monograph's subject matter, and even more so, on where the monograph is housed. Factors such as the age of the monograph and how often it is checked out also influence the likelihood of being missing, but only in a minor way.

With regard to the condition of monographs at the EAST libraries, there is significant variation based on age and frequency of use, but again the subject area of the monograph and where it is housed proved to be even more significant,

though that may be due to differences across evaluators. Suggestions to future cohorts or others undertaking retention projects is to consider retaining additional copies of older materials, along with additional copies of items at libraries known to have inventory issues, or to validate retention copies in those locations.

Opportunities for further work around validation and risk assessment abound. EAST was contacted by Adam Chandler of Cornell University Library in late 2017 to ask if any data had been collected regarding the use of security systems (e.g., Tattle Tape and rfid systems) and its correlation to missingness. EAST found this to be an interesting question though had no immediate data. A quick survey to EAST members was compiled and administered, and while the results are complex, they did not immediately reveal large differences in missingness rates between low and high security environments. Cornell plans to replicate the EAST validation study and to study loss rates and may have more to say on this topic in the future.

Areas where EAST could continue its validation work include potentially validating the approximately 9,000 titles which were statistically determined to be at risk for which no surplus copies existed in either cohort. Also revalidating titles over time could help to determine loss rates which might inform future projects. EAST has not explored journal validation at other than the volume level, nor done any comparisons with digitized volumes. EAST is in talks with BookTraces (<http://www.booktraces.org/>) and exploring ways in which they might collaborate on determining the risk of losing monographs with unique artifactual attributes.

In April of 2018, with funding from the Mellon Foundation grant, EAST sponsored a summit on monograph shared print, inviting representatives from the major shared print monograph programs in North America as well as thought leaders in shared print and related topics such

as preservation and digitization. One of the outcomes from this summit was the formation of a working group to look at standardized methods of assessing risk and condition of the collective collections, potentially following up on some of the work conducted on optimal number of copies of JSTOR journals.⁹ As discussed above, the subjective nature of condition assessment used by the EAST validation study proved problematic. To quote an unpublished report titled *Regional Differences in Library Material Conditions*, "To suppose that keeping only a random handful of copies will be sufficient to preserve that work is folly." Factors such as the physical and chemical condition of the paper, historic and future storage conditions, age of the material, and perhaps external factors such as security, location, and type of the holding institution may all factor into risk assessments and determining the number of usable copies available now and in the future. EAST welcomes and looks forward to opportunities to be involved in future research in these areas of validation and risk assessment.

References

- Garabedian, Michael. "Condition Considerations: An Inquiry Into Recording Conditions in Consortial Collections for the Purpose of Selecting (and Deselecting) Shared Print Copies." Proceedings of the Charleston Library Conference, November 2014. <http://dx.doi.org/10.5703/1288284315571>.
- Hale, Dawn. *Shared Collections: Collaborative Stewardship*. Chicago: ALA Editions, 2016.
- Koch, Teresa, and Andrew J. Welch. "Monograph Validation Strategies in Shared Print Programs: Variations and Value." *Collaborative Librarianship* 8, issue 3, article 7 (2016). <https://digitalcommons.du.edu/collaborativelibrarianship/vol8/iss3/7>.
- Koch, Teresa, and Andrew J. Welch. "Validation of Shared Print Collections. Central Iowa Collaborative Collections Initiative Survey." Presentation. https://ci-cci.org/files/2014/06/ala2017_pan_rev062017-2.pdf.
- Nadal, Jacob, Annie Peterson, and Dawn Aveline. "Scarce and Endangered Works: Using Network-Level Holdings Data in Preservation Decision-Making and Stewardship of the Printed Record." <http://www.jacob-nadal.com/wp-content/uploads/2011/05/ScarceAndEndangeredWorksv7.pdf>.
- OCLC. "Weeding and Deselection Bibliography." Sustainable Collection Services. <https://www.oclc.org/en/sustainable-collections/bibliography.html>
- Schonfeld, Roger C., and Ross Housewright. "What to Withdraw: Print Collections Management in the Wake of Digitization." *Ithaka S+R*. Last Modified September 1, 2009. <https://doi.org/10.18665/sr.22357>.
- Stearns, Susan. "EAST by Northeast." *Collaborative Librarianship* 8, issue 1, article 4 (2016). <https://digitalcommons.du.edu/collaborativelibrarianship/vol8/iss1/4>
- Yano, Candace Arai, Zuo-Jun Max Shen, and Stephen Chan. "Optimising the Number of Copies and Storage Protocols for Print Preservation of Research Journals." *International Journal of Production Research* 51, issue 23-24 (August 2013): 7456-7469. <https://doi-10.1080/00207543.2013.827810>.



Appendix

This document is available on the EAST website:

https://eastlibraries.org/sites/default/files/BLC_Uploads/ValidationResults-AppendixA.pdf

Endnotes

¹ WEST Standards for Issue and Volume Level Validation, available at: https://www.cdlib.org/services/west/docs/WESTStandards_Issue_VolumeLevelValidation.docx.

² Teresa Koch and Andrew J. Welch, "Monograph Validation Strategies in Shared Print Programs: Variations and Value," *Collaborative Librarianship* 8, issue 3, article 7 (2016), <https://digitalcommons.du.edu/collaborativelibrarianship/vol8/iss3/7>.

³ EAST validation study accessing condition video: <https://sites.google.com/a/blc.org/validation-study-coordination/training-materials/assessing-condition>

⁴ Google App Script: <https://developers.google.com/apps-script/>

⁵ EAST validation tool code: <https://github.com/samato88/EastValidation-Tool>

⁶ Koch and Welch.

⁷ Koch and Welch.

⁸ The EAST Retention model is described more fully at: <https://eastlibraries.org/data-results>.

⁹ Candace Arai Yano, Zuo-Jun Max Shen, and Stephen Chan, "Optimising the Number of Copies and Storage Protocols for Print Preservation of Research Journals," *International Journal of Production Research* 51, issue 23-24 (August 2013): 7456-7469, <https://doi.org/10.1080/00207543.2013.827810>.