2018 Sample Inventory of CUL Open Stacks: Where are our Books?

29 November 2018 10:35

## Investigators

* Adam Chandler, study design, data analysis
* Wendy Wilcox, study design, project management and data collection
* Joanne Leary, study design, data collection

## Acknowledgements

* Jenn Colt, Google App installation
* Sara Amato (EAST consortium data analyst) assisted with sampling method, software installation, and data sharing
* Darla Critchfield, Kim Laine, Michelle Hubbell, Bethany Silfer Law Library and Mann Library Access Services staff, data collection

## Introductions

The impetus for conducting this validation study stemmed from a 2017 Taskforce Report on the elimination of tattletaping for open stack materials. Despite the charge to discontinue tattletaping of library materials, members of the group could not recommend proceeding with the plan given significant opposition to this decision. “Even though members of our task force are not confident that it is actually preventing theft we must recommend continuing tattle-taping because staff, primarily selectors, clearly oppose changing the policy at this time.” The task force felt that quantitative data was necessary before proceeding and recommended conducting an inventory of the library’s open stacks collections using the methodology (and tools, perhaps) employed in the EAST validation study to use as a baseline to inform present and future decision making on this issue.

The Eastern Academic Scholar’s Trust is a partnership of college and university libraries dedicated to the shared retention of print resources. Fifty-two libraries are retention partners in the consortium; partners range from small liberal arts institutions such as Bryn Mawr and Smith College to larger universities such as New York University, Florida State University and University of Pittsburg. As part of their retention agreements, the libraries agreed to participate in a validation study designed to quantify the likelihood of finding a monograph in each institutions library stacks. “In order to evaluate the statistical likelihood that a retained volume exists on the shelves of any of the institutions, the EAST incorporated sample-based validation studies. The specific goals of this study were to establish and document the degree of confidence, and the possibility of error, in any EAST committed title being available for circulation. Results of the validation sample studies help predict the likelihood that titles selected for retention actually exist and can be located in the collection of a Retention Partner, and are in useable condition.” [<https://eastlibraries.org/validation>]

Overall, EAST reported a 97% availability rate. The aggregated results from both cohorts (312,000 holdings across the 52 libraries) showed:  
\* 97% of monographs in the sample were accounted for: mean: 97%, median: 97.1%, high of 99.8% and low of 91%. (Note: “accounted for” includes those items previously determined to be in circulation based on an automated check of the libraries’ ILS)  
\* 2.3% of titles were in circulation at the time of the study  
\* 90% of the titles were deemed to be in average or excellent condition with 10% marked as in poor condition. Not surprisingly, older titles were in poorer condition.  
\* The most significant factor for an item being missing was the holding library.  
The Access Services Committee and Adam Chandler, chair of the Tattletape Task Force, were charged with conducting a validation study of Cornell University Library’s open stacks. Given the extensive data provided by the EAST validation study, it was decided that CUL would sample our collection using their study methodology. Our goal was to benchmark our collection against the 52 partner libraries in EAST as a means of understanding the availability of the CUL collection.

## Data collection

We are grateful to the East Consortium for generously sharing their methodology and associated Google App for data collection. Jenn Colt was able to take their models and create a Cornell instance of the Google App for the Cornell Validation Study.

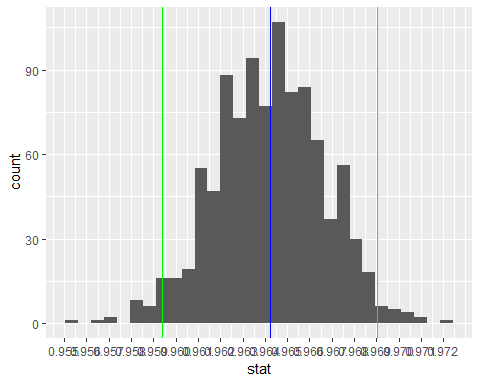
We sampled 6006 monograph across campus. Some caveats about Cornell’s data sample: the Library Annex was excluded because the stacks there are closed; Fine Arts was excluded because they are in the middle of a building transition; location codes in HLM were merged into one group. Finally, some items were removed from the dataset after the initial sample because they did not actually fit the criteria of our study (no circ items), leaving `r nrow(df) items in our dataset. (see Appendix A).

Wendy Wilcox led the team that did the data collection in the stacks which too place between April and July 2018. The data collection team worked through the sample dataset, verifying the presence of items in the stacks and the corresponding condition of the item. For items not found, the team checked item statuses in the library catalog. Moving forward in our report, AF (accounted for) will equal checked out items plus items verified as present in the stacks.  
AF (accounted for) = checkedout + present

Cornell’s aggregate AF rate is 96.4% and we are 95% confident that the true proportion of accounted for monographs across CUL is between 0.959, 0.969 (see Appendix B).

AF (accounted for) = checkedout + present

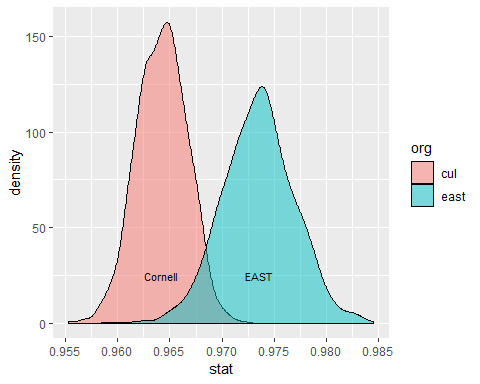
## Bootstrap simulation to estimate standard error



Cornell’s aggregate accounted for rate is 96.4% and we are 95% confident that the true proportion of accounted for monographs across CUL is between 0.959, 0.969.

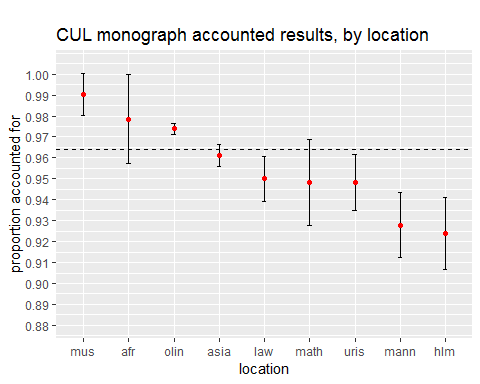
## How does Cornell compare to the EAST consortium libraries?

# combine CUL and EAST into one boot df for visualzation  
  
  
t\_boot <- boot  
t\_boot\_east <- boot\_east  
  
t\_boot$org <- "cul"   
t\_boot\_east$org <- "east"  
  
boot\_all <- rbind(t\_boot, t\_boot\_east)  
  
ggplot(boot\_all, aes(x = stat, fill = org)) +  
 geom\_density(alpha=0.5) +  
 annotate("text", x = p\_hat, y = 25, label = "Cornell", size = 3) +  
 annotate("text", x = p\_hat\_east, y = 25, label = "EAST", size = 3) +  
 #scale\_x\_continuous(limits = c(.95, .99), breaks = seq(.95,.99, by = .1))  
 #scale\_x\_continuous(labels = scales::percent %>% round(1))  
 scale\_x\_continuous(breaks = c(.955, .960, .965, .970, .975, .980, .985, .990))



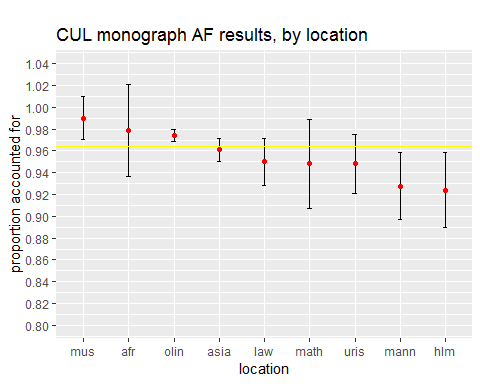
## There significant differences across CUL units in the percentage of monographs accounted for.

## # A tibble: 9 x 5  
## location\_group mean se total num\_missing  
## <fct> <dbl> <dbl> <int> <dbl>  
## 1 mus 0.990 0.0100 102 1  
## 2 afr 0.978 0.0212 49 1  
## 3 olin 0.974 0.00276 3221 84  
## 4 asia 0.961 0.00530 1282 50  
## 5 law 0.950 0.0107 400 20  
## 6 math 0.948 0.0205 116 6  
## 7 uris 0.948 0.0134 270 14  
## 8 mann 0.928 0.0155 280 20  
## 9 hlm 0.924 0.0172 210 16



## Tattle-tape: What evidence do we have that security stripping improves AF rates? In other words, what is our estimate of the effect size (i.e., return on investment) for libraries that operate security systems?

### Cornell



At Cornell, we had an experiment ready to be conducted because there is one unit that does not use security stripping or gates, Law. Our intuition might tell us that the Law AF rate should therefore be lower than the other units. That is not the case. The Law mean AF (accounted for rate) in this sample is right in the middle of pack, with confidence intervals overlapping other units that have both higher and lower AF rates.

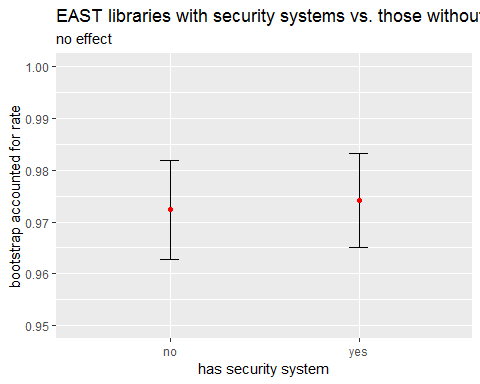
### EAST surveyed libraries that participated in their validation study about security practices

Sometime after completing it’s validation study, EAST conducted a survey of participating libraries to find out about the theft deterrence practices. 32 libraries responded. The library names are anonymized.

|  |  |  |
| --- | --- | --- |
| library | tattletape\_yes\_no | validation\_score |
| anteater | Yes | 0.984 |
| armadillo | No | 0.948 |
| axolotl | No | 0.984 |
| buffalo | Yes | 0.935 |
| camel | Yes | 0.976 |
| chameleon | Yes | 0.951 |
| cheetah | No | 0.990 |
| chipmunk | Yes | 0.953 |
| chupacabra | No | 0.973 |
| crow | Yes | 0.975 |
| dolphin | Yes | 0.989 |
| giraffe | Yes | 0.994 |
| grizzly | Yes | 0.990 |
| hedgehog | No | 0.953 |
| hippo | Yes | 0.963 |
| ifrit | Yes | 0.978 |
| iguana | No | 0.956 |
| jackal | Yes | 0.997 |
| koala | No | 0.983 |
| lemur | Yes | 0.916 |
| leopard | Yes | 0.992 |
| liger | Yes | 0.988 |
| llama | Yes | 0.982 |
| manatee | No | 0.984 |
| monkey | Yes | 0.953 |
| narwhal | No | 0.970 |
| nyan cat | Yes | 0.994 |
| otter | Yes | 0.990 |
| panda | Yes | 0.968 |
| quagga | No | 0.982 |
| squirrel | Yes | 0.995 |
| wombat | Yes | 0.967 |

## # A tibble: 2 x 2  
## tattletape\_yes\_no mean  
## <fct> <dbl>  
## 1 No 0.972  
## 2 Yes 0.974

## # A tibble: 2 x 5  
## af se total status n   
## <dbl> <dbl> <int> <chr> <chr>  
## 1 0.972 0.00476 1000 no 10   
## 2 0.974 0.00455 1000 yes 22



In this experiment we divided the EAST libraries into two groups, the 22 libraries in the survey with security systems vs. the 10 libraries with no security systems, and generated accounted for (AF) rates and standard errors using bootstrap simulation for each group. The difference in AF rates can be explained by random noise, as we see from the overlapping 95% confidence intervals. We conclude again that the effect size of having a security system is zero.

## Conclusions

The validation study was useful in determining the availability of CUL materials in the open stacks; when patrons seek a monograph in the stacks, they are able to locate that item 96.4% of the time. This translates into a loss rate of 3.6% of materials in the open stacks.

[Not sure about this paragraph. It seems to contradict our conclusion. - Adam] Tattletaping prevents theft of materials. However, we know that the theft of materials is only one component of lost or missing items. We know from conducting tracers for wanted items, that a percentage of lost materials are actually due to misfiling of materials in the open stacks.

We wish to decouple the idea that tattletapping is directly correlated to reducing lost items in the open stacks. As we have shown, there is minimal differences in the AF rate in libraries that tattletape versus libraries that do not tattletape.

## Recommendations

1. Where confidence intervals are widest, do more sampling in Cornell unit libraries to improve the accuracy of our estimates.
2. Security stripping is only one variable that might account for differences in the percentage of items accounted for across different libraries, and from the empirical evidence we’ve gathered, it isn’t a variable with any predictive power here or across the EAST consortium. Therefore, we believe that a better investment is more support for open stacks control and management at units with lower accounted for rates.
3. Create a statistical model that includes multiple variables to identify items with higher probability of being unaccounted for in the open stacks. A statistical model will inform us about the factors at play and help us to increase the availability of open stacks titles. This work is already underway.

# Appendix

INSERT Adam and Joanne’s explanation about the sample parameters.