Abstract
This study explores student attitudes on the library’s role in their success, with a focus on library building use, library resource use, library instruction, and how student attitudes and library use behaviors may differ by discipline. Quantifying the relationship between library usage and student success is one of the six areas recommended for further research by ACRL. This study focuses on the variation and the uncertainty of the measurement of this relationship across disciplines, using Bayesian multilevel regression methods.

Levels of library resource usage and percentages of those believing the library contributes very much to the respondents’ academic success vary quite a bit by discipline. Cumulative GPA is higher on average for students who use library resources more frequently, but not for students who receive library instruction, in this study. Although for undergraduates, higher frequency of building usage predicted higher probability of believing the library contributes very much to academic success, it did not predict higher GPA.

Introduction
Quantifying the relationship between library usage and student success is one of the six areas recommended for further research in the 2017 ACRL report, “Academic Library Impact: Improving Practice and Essential Areas to Research.” This study focuses on the variation and the uncertainty of the measurement of this relationship across disciplines. Describing the relationship as it varies on a group level is accomplished by multilevel regression modeling, which combines a model for the individuals with a model for the groups.

The group-level information derived from this study can be used for two purposes: mapping the correlation between library instruction, library resource use and positive attitudes and cumulative GPA; and communicating the value of library instruction and resource use to university administration and faculty, who have observed the relationship between the library and their students within their own discipline, but perhaps not so much for other disciplines. It may be more convincing to see results that reflect their own experience, and it is more informative to study the variations than to limit the results to an average.

Although the results from most published work about the impact of library usage and instruction center around statistical significance, this paper follows the recommendations of the American Statistical Association (Wasserstein and Lazar) to present estimates with measures of uncertainty and of McShane, Gal, Gelman, Robert, and Tackett to “abandon statistical significance.” Statistical significance calculations are based on a null hypothesis; in real life social science, treatment effects or differences between groups are highly unlikely to be zero. They are likely to be small and varying between individuals, groups, and situations. Dividing results into categories of significant or non-significant by the arbitrary .05 p-value level often leads to a misunderstanding that a p-value greater than .05 is proof that there is no treatment effect. On the other hand, too much confidence is placed in low p-values, as studies with low p-value results often fail to replicate. Recommendations include presenting estimates
with Bayesian uncertainty or credible intervals, which present the probability that the value is really within the interval, given the data, with the uncertainty of the estimate.

**Literature Review**

Libraries have long been interested in analyzing user behaviors and attitudes toward library use. The idea of measuring library service impact as a means of communicating library value to stakeholders gained significant ground in 2010 through the release of ACRL’s Value of Academic Libraries report. That report recommended that academic libraries seek to track “library influences on increased student achievement,” and specifically to seek ways to link ostensibly objective indicators of success, such as student grade point averages and test scores, with their library usage experiences. In addition, the report noted heavy investments from academic libraries in library instruction to students but a general lack of program-wide and longitudinal assessments that measured real learning resulting from those instruction sessions, that might collectively in turn communicate library value more concretely to external stakeholders. An additional area of interest from the report focused on seeking to measure students’ experiences with the library, including their “attitudes, and perceptions of quality” through the collection of data.

Subsequent to the publication of the Value of Academic Libraries Report, an increasing number of library service impact studies began to emerge. In one early study, Wong and Webb (2011) sought empirical evidence of the positive impact of library service usage on student GPA. The researchers conducted a massive study of over 8,500 students at Hong Kong Baptist University that analyzed students’ graduation GPA and the number of times these students had checked out material from the library. Results were organized by academic discipline and student rank (undergraduate or graduate), with the finding that 65% of the academic subject areas considered showed positive correlations with library checkouts and higher graduation GPA. Subsequent research (Soria, Fransen, and Nackerud, 2013, Stone & Ramsden, 2013) has also concluded that book checkout, library database usage, and e-journal access/usage have positive correlations with student GPA and other forms of academic achievement. It is noteworthy that these library-use indicators are all student self-directed library research activities that require engagement, activity, and often some sort of procedural follow-through from the student.

Other library value studies have often focused on the relationship of library instruction with student success, but with varying results. Wong and Cmor (2011) investigated the relationship of attending discipline-relevant library instruction workshops and student graduation GPA, using data from a previous study (Wong and Webb, 2011) of 8,500 students at Hong Kong Baptist University. Overall, this research found that results varied by student academic disciplines and by number of instruction workshops attended. Data from students in only 24.4% of the subject area groups considered showed positive correlations (at varying levels of significance) with a higher graduation GPA. The researchers also found that attending more sessions of discipline-relevant library instruction increased the likelihood of a stronger positive correlation with higher GPA at graduation.

In a study analyzing 4,489 transcripts, Bowles-Terry (2012) found a very small yet statistically significant positive difference in GPA among students who had been enrolled in an upper-level course that received course-related library instruction (CRI) and whose attendance at that instruction session was inferred, while Soria, Fransen, and Nackerud (2013) found probable CRI attendance by 5,368 first-year undergraduates was correlated with a lower GPA; little explanation was given for this outcome other than noting the difficulties of assessing instruction effectiveness. However, a subsequent study by
Soria, et al. (2017)\textsuperscript{13} used different analysis methods and found opposite results, that “students who engaged in libraries instruction” by attending library workshops, CRI, or a library class had a significantly higher GPA than did those students who had not been involved in library instruction.

Gaha, Hinnefeld, and Pellegrino (2018)\textsuperscript{14} sought to normalize data to account for probable disciplinary grading variations (a multilevel model would make this step unnecessary, even with a small dataset). While they found some positive correlations between higher normalized GPA among students who had taken at least one course with library instruction, again there was no control for whether a student had actually attended that instruction session, so conclusions are somewhat speculative. Krieb (2018)\textsuperscript{15} analyzed the impact of CRI and reference desk usage on student retention and GPA at a community college, finding some positive impact of CRI completion on retention but minimal impact on GPA.

Many big data library value studies have relied heavily on identifying correlations and significance to show which library activities, behaviors, resources or service usage seem to matter most in terms of supposedly objective student success variables such as GPA. Most of the analyses accounted for the average differences in GPA across majors or colleges. Soria, et al. (2017)\textsuperscript{16} did so by allowing varying intercepts for each college, but did not allow for varying relationships with the library (different regression slopes for the library usage variables).

Wong and Webb\textsuperscript{17} and Wong and Cmor\textsuperscript{18} realized that not only the average GPA varies by major, but that different majors have different relationships with the library and for that reason conducted analyses (Pearson correlation and chi-square tests, respectively) for each discipline group separately. Because some discipline groups had few students, these groups were left out of the analysis. A multilevel regression could have combined both of these analyses into one, while allowing use of data from all discipline groups and using the number of library sessions per discipline as a group-level predictor.

Another challenge faced in a number of big data instruction-value studies is how to determine whether a student has indeed received library instruction. One common method has been to cross-check the institution’s student course enrollment data with library instruction program records of CRI sessions for that course (Bowles-Terry\textsuperscript{19}; Gaha, et al.\textsuperscript{20} Soria, et al.\textsuperscript{21} ), while other studies (Wong and Cmor\textsuperscript{22}) may be able to verify completion of an initial mandatory library orientation session but not attendance at subsequent and potentially more relevant CRI sessions. These methods do not guarantee the student was actually present at a CRI session. In addition, some researchers state that instruction program records may not be complete, and are particularly problematic for high enrollment multi-section courses for which CRI is offered for some but not all sections. Krieb\textsuperscript{23} described an alternate method of asking students after CRI sessions for their verbal consent to be tracked for research purposes. This method allowed solid identification of both the course and CRI completion, as well as close tracking of individual student retention and GPA success measures. While this method ensures a student did receive CRI, larger college and university libraries may find it difficult to replicate the individual student consent process for reasons of scale.

**Methods**

A survey including undergraduate and graduate students was conducted in 2016 for the purpose of service improvement. The survey sample was stratified on college and student type (undergraduate or graduate) combination. Within each stratum, a systematic sample was drawn, sorting by gender and...
residency (Iowa, non-Iowa US, International), and selecting each nth student to reach the specified percentage for that stratum. All regression models included the college and student type sampling strata. Survey sample size and response rates are given in Table 1. Details of the statistical analyses are in the Appendix.

Table 1. Survey sample size and response rates.

<table>
<thead>
<tr>
<th></th>
<th>Undergraduate Students</th>
<th>Graduate Students</th>
<th>Postdocs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eligible Sample</td>
<td>2952</td>
<td>3451</td>
<td>318</td>
</tr>
<tr>
<td>Completed Surveys</td>
<td>494</td>
<td>1116</td>
<td>117</td>
</tr>
<tr>
<td>RESPONSE RATE</td>
<td>16.7%</td>
<td>32.3%</td>
<td>36.8%</td>
</tr>
</tbody>
</table>

Three outcome variables were modeled:

1. Students’ library resource usage index (LRU) calculated from responses to survey questions about frequency of use of various library resources and services
2. Students’ perception of library contribution to academic success (LCS)
3. Students’ cumulative GPA

**Library resource usage index (LRU)**

The LRU was constructed from survey questions (see Appendix) about the frequency of use of the library website, the library search, article indexes and databases, electronic journals and articles, e-books, physical items (e.g., books, DVDs, CDs, maps, microforms, etc.), and of materials from Special Collections and University Archives.

Each of these survey questions had response categories of daily, weekly, monthly, once a semester, less often, and never. The LRU, simulating an estimate of a total average monthly library usage count was created by setting a response of daily to 16 times a month; weekly to 4 times a month; monthly to 1 time per month; once a semester to 0.3 times per month; less often to 0.2 times per month; and never to 0. These estimated counts were summed over the set of questions. If a student did not answer at least five of these questions, the case was deleted. Sixteen cases were deleted, leaving 1595 cases for analysis.

Evidence that this is a fairly reliable way to estimate the frequency of usage comes from a previous analysis where an estimate of graduate and undergraduate website usage derived in this way was similar to an estimate of undergraduate and graduate website usage derived from a time series regression of web log data (Anderson). The LRU should be regarded as an index or artificial construct. Summing the estimates from the individual resource questions may inflate an estimate of total resource usage; for instance, the respondent may use the library website to find a database, which is then used to find an article, counting the use three times.

**Perception of library contribution to success (LCS)**

The LCS variable is the binary outcome variable constructed from the response to the question “To what extent have the Library’s resources and services contributed to your academic success?” A response of “very much” was set to 1, and missing or other responses (“some”, “very little”, or “not at all”) were set...
to 0. Most responses were either “very much” or “some”. Making it a binary variable was a way to simplify the analysis. Since not many answered very little or not at all, we focused on the respondents who answered “very much.”

**Cumulative GPA**

There are fewer observations (1551) for the cumulative GPA analyses than for the other analyses because the GPA data was requested retroactively in 2018 and not all cases were matched. The average of the GPAs of the undergraduate survey respondents is 3.05, while the mean for all undergraduates at this institution is 3.03.

Because the range of the GPAs is bounded by 0 and 4 and the distribution is extremely left-skewed, a beta regression with a logit link was used to model the GPAs (Smithson and Verkuilen). A beta distribution is a distribution of probabilities between 0 and 1 and can take many shapes, from symmetrical to skewed. The cumulative GPA variable was divided by 4 and transformed so that values were between 0 and 1 and not exactly 0 or 1. The beta regression fits much better and does not predict impossibly high values (Figure 1). The logit link means that instead of the GPA itself being modeled, a function of the GPA (the logit function) is modeled instead.

![Distributions of observed data and a random sample of replications](image1)

Figure 1. The upper right histogram displays the extreme left-skewness of the cumulative GPA distribution, which is bounded at the upper end at 4.0. The two plots on the left compare the densities of replications (light blue lines) from a linear regression and a beta regression with the distribution of the observations (dark blue line). The replications from the linear regression (top) follow a symmetrical (normal distribution) distribution matching the observations very poorly, while the replications from the beta regression match the observations fairly well.
Multilevel regression

In a linear regression, each regression coefficient (slope) may be interpreted as the predicted difference in the outcome variable when the predictor variable is one unit greater (or for categorical variables, present or not present) while the value of all other predictors remain the same. Predictive does not imply causal, but simply compares the averages of the outcomes for different groups, given different conditions (Gelman and Hill)\textsuperscript{27}.

Respondents from various demographic, discipline, class level, and graduate students with and without thesis/dissertation requirements may respond differently to questions about the library and have varying relationships with the library. The variation arising from these groups needs to be accounted for in the model so that the importance of other predictors may be compared more directly (McElreath)\textsuperscript{28}.

A common method is to include a dummy variable equal to 1 if a respondent is a member of a particular group and equal to 0 if not. This allows estimates of a different mean or regression intercept for each group. The intercept for each group is estimated separately, so for a group with few members, the estimate may be unreliable. The result is a set of parallel regression lines (same slope but different intercepts).

An alternative is a multilevel model, where separate intercepts can be estimated using all the data. Where there are fewer members, the group mean will be closer to the overall mean. This is called partial pooling, and offers a conservative estimate of the differences between groups, alleviating the noisiness and unreliability of small group sample sizes. In addition to differing means of the outcome variable for different groups, there may be different slopes for predictor variables within each group. Although these different slopes for different groups (called interactions) can also be estimated with the dummy variable method, again there is the advantage of partial pooling offering a more conservative and reliable estimate (Gelman and Hill)\textsuperscript{29}. McElreath\textsuperscript{30} argues that multilevel regression should become the default method of regression, to account for differences between individuals and groups that can’t be measured.

Three possible sets of groups were available for use in the multilevel models. The first, college and graduate/undergraduate type combination, were the sampling strata for the survey and are included to account for the varying probability of selection.

The second group, discipline and graduate/undergraduate type combination, overlaps with college and type. Some colleges at the authors’ institution are represented by one discipline, other colleges contain multiple disciplines whose library usage and attitudes toward library use may differ significantly, and some disciplines are found in multiple colleges.

Disciplines may be a more precise grouping to describe this varying relationship with the library than are colleges. In addition, disciplines make it easier to compare our results to results from other universities, where the college organization may be quite different. The number of respondents in each discipline/type group are given in Table 2.
Table 2. Number of respondents per discipline.

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Graduate</th>
<th>Undergraduate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag &amp; Life Sciences</td>
<td>123</td>
<td>121</td>
</tr>
<tr>
<td>Arts &amp; Design</td>
<td>58</td>
<td>63</td>
</tr>
<tr>
<td>Business &amp; Econ</td>
<td>77</td>
<td>50</td>
</tr>
<tr>
<td>Humanities</td>
<td>30</td>
<td>21</td>
</tr>
<tr>
<td>Math &amp; Comp Sci</td>
<td>54</td>
<td>12</td>
</tr>
<tr>
<td>Phys Sci &amp; Engineering</td>
<td>393</td>
<td>117</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>155</td>
<td>99</td>
</tr>
<tr>
<td>Interdisciplinary</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Vet Med</td>
<td>219</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1116</td>
<td>494</td>
</tr>
</tbody>
</table>

The level of the student (classification 1, 2, 3, and 4 for undergraduates, and master’s degree with no thesis, master’s degree with thesis, and PhD for graduate students) is the third overlapping grouping.

Bayesian methods were used in this study as an alternative to the conventional null hypothesis significance testing and to present ranges of values for parameter estimates, which can be used to estimate the probability that an effect is greater or less than zero, or greater or less than the effect for another group, given the data.

Bayesian estimation starts with a “prior” belief about the values of the parameters to be estimated. These priors may come from previous research. The data are then used to update that prior estimate. The result is a distribution of plausible estimates, given the data, called a posterior distribution. In this study, we assume we know nothing and do not use previous research to set priors. The priors used in this analysis are distributions centered around zero, called weakly informative priors31, which allow the data to have most of the influence while restraining unrealistic estimates and stabilizing the computation. Again, this produces conservative estimates since the starting point is the area near zero, while leaving room to be a reasonable distance from zero as well. The most plausible values of the parameters we are trying to estimate, given the data, are the most frequent values in the posterior distribution.

Newer advanced software, such as the rstanarm R package32 used in this study, and more computing power allow more computationally stable estimation of these complex models than were possible using non-Bayesian methods. Muth, Oravecz, and Gabry33 provide a tutorial for using rstanarm and a brief review of Bayesian data analysis, with a more complete explanation of Bayesian methods from Gelman and Hill34, Kruschke35, and McElreath36.
Multilevel regression and poststratification (MRP)

To estimate average levels of the LRU and of the LCS within each discipline and type, multilevel regression and poststratification (MRP) was used to adjust for differential nonresponse rates by gender and international status. The multilevel model estimates the average for each college, discipline, type, and demographic combination, which is then weighted by the population for each combination (Gelman and Hill)\textsuperscript{37}. Females made up 43% of the university student population and 54% of the response, while international students made up 12% of the university student population and 31% of the response.

Average predictive comparisons

Each of our three outcome variables were modeled with non-linear models and interactions between variables, making the direct interpretation of regression coefficients difficult, unlike in linear regression. In these situations, average predictive comparisons allow intuitive comparisons to be made.

To make average predictive comparisons, the values at which to compare an input are chosen, the input variable is set to the first value over all observations, predictions made, then the input value is set to the second value overall and predictions made again. The average difference between the set of predictions at the two different values is the average predictive comparison. For two-way interactions, the values of both inputs can be changed to clearly show the predicted differences (Gelman and Hill\textsuperscript{38}, Gelman and Pardoe\textsuperscript{39}).

Graphical display of results

Results of the models are communicated primarily in graphical form. This allows the easiest comparison of the size of the effects of the predictors between each other and across groups, and also communicates the amount of uncertainty in the estimates at a glance. Graphs are constructed in the R package ggplot2\textsuperscript{40}.

Average predictive comparison plots, such as Figures 3 and 4, show the size of the predicted difference in the outcome as the value of one predictor is changed, for each predictor listed on the y-axis. The median of the plausible values for the predicted difference is presented as a white dot. The thick gray bars represent the 50% credible interval (the middle 50% of the plausible values from the posterior distribution). There is a 50% chance that the unobserved parameter is within that range, given the data. The thin lines represent the 90% credible interval. The longer the bars and thin lines, the greater the uncertainty of the estimate.

A vertical dashed line is drawn at zero on the x-axis. A dot and/or bar further to the right of the zero line than other dots or bars means that that predictor predicts a greater positive difference in the outcome, by itself, than other predictors do. If the entire 50% interval is to the right of zero, then there is at least a 75% probability that the true value is greater than zero. If the entire 90% interval is to the right of zero, there is at least a 95% probability that it is greater than zero.
Results

Levels of library resource usage (LRU) and percentages of those believing the library contributes very much to the respondents’ academic success (LCS) varied quite a bit by discipline. Both were quite a bit higher for graduate students than for undergraduates. LRU and LCS tended to be high in the same groups, with Humanities the highest in usage and nearly the highest in success (Figure 2).

![Library Resource Usage Index (LRU) by discipline](image)

![Library contributed to your academic success? By discipline](image)

Figure 2. Average levels of the LRU (sum of estimated uses per month) by discipline for graduate and undergraduate students (left) and the proportion who believe that the library’s resources and usage contribute “very much” to their academic success (LCS) (right). Wider bars indicate more uncertainty or a wider range of plausible values for the estimate, caused by fewer respondents in the group or a wider spread in responses. The thicker segments represent the middle 50% of the plausible values. Partial pooling shrinks the estimates toward the mean, especially for groups with few respondents.

Reporting library instruction in a class predicted higher LRU (Figure 3). For graduate students, LRU was predicted to be 15% higher for US residents reporting CRI and even higher for international students at 39%. The difference in LRU for undergraduates reporting CRI was larger, with estimates of 64% for US residents and 39% for international students, but there was no clear difference between US residents and international students as the 50% credible intervals overlap.
Figure 3. The perception that the library contributes very much to a graduate student’s success predicts about a 115% higher LRU, compared to a graduate student who does not have this perception. For undergraduates, it predicts about 77% higher LRU. CRI predicts about 25% higher LRU for graduate students and 50% for undergraduates. LRU of international students (without CRI) is higher than U.S. residents. LRU of female graduate students is about 13% lower than that of male graduate students, while LRU of female undergraduates is about 13% higher than that of male undergraduates.

In Figure 4, the average difference in LRU between those who believe the library contributes very much to success and those who do not varies quite a bit by discipline, but it is always positive. Veterinary Medicine students with LCS average 200% higher LRU than those without, while Physical Sciences & Engineering, Math and Computer Science, and Humanities graduate students with LCS have about 75% higher LRU than those without. Among undergraduates, Physical Sciences & Engineering is one of the disciplines with the largest differences (150% higher LRU with LCS), and Business & Economics among the lowest (around 40% higher LRU). In contrast, for CRI, there is less variation across disciplines in the difference in LRU.
International students tended to report higher LRU than US residents and were also about 6% (undergraduates) to 8% (graduates) more likely to say the library contributes “very much” to their academic success, for both graduate and undergraduate students. Female graduate students were about 9% more likely than male graduate students (7% for undergraduates) to say the library contributes “very much” to their success (Figure 5).

US residents who had received course-related instruction had a higher predicted probability of LCS than US residents who had not received CRI. For international students, receiving CRI made virtually no difference, after adjusting for LRU, keeping in mind that CRI was associated with a bigger increase in LRU for international graduate students (Figures 3 and 5).
Building usage was a good predictor of undergraduates’ LCS (11% higher predicted probability, comparing 4 times per month to once per month), which was more than LRU predicts. For graduate students, building usage predicted only a slight difference in LCS. (Figure 5).

For most undergraduate disciplines, both higher LRU and higher building usage predicted a greater probability of LCS, a little over 10% for building usage and a little under 10% for LRU. Humanities had a larger difference, about 12% for LRU and 17% for building usage, comparing 4 times per month to 1 time per month (Figure 6.)

For graduate students, LRU predicted slightly higher probability of LCS for international students than for US residents. The predicted difference in probability for LCS with higher LRU ranged from about 20-25% for Interdisciplinary, Humanities, and Arts & Design down to 10% for Business & Economics and Mathematics & Computer Science graduate students (Figure 6.)
Figure 6. Subsetting by each discipline and comparing the predicted probability of saying that the library contributes “very much” to the respondents’ success, at a value of 4 times per month and at once per month, for LRU and for building usage. Squares indicate US residents and triangles indicate international students, which shows a little higher LCS with higher LRU for international graduate students. Otherwise the two groups are very close.
Model of Cumulative GPA

Cumulative GPA was higher on average for students who used library resources more frequently, but not for undergraduate students who received library instruction. An undergraduate with an LRU of 4 times per month averaged a higher GPA by 0.1 than an undergraduate with an LRU of 1, if they had not received CRI. The difference was greater than zero with at least 95% probability, given the data. Little to no difference in GPA was predicted for those undergraduates who reported CRI and higher LRU.

The average predicted graduate student GPA was higher by about 0.05, at an LRU of 4 times per month vs. 1 time per month. This was true whether or not the student reported receiving library instruction. The difference was greater than zero with at least 95% probability. For international students and US residents, the average predicted difference in GPA at the two levels of LRU was the same for graduate students and very close for undergraduates (Figure 7).

On average overall, after adjusting for LRU, both graduates and undergraduates who had received CRI had the same or very slightly lower GPAs than those who had not. There was some variation across disciplines, notably, the effect was negative for Humanities undergraduates (-0.1) and Ag & Life Sciences undergraduates with at least 75% probability. There was no difference between international and U.S. residents, which was true across disciplines (fourth line of Figure 7 and second column of Figure 8).

![Average predictive comparisons for cumulative GPA](image)

*Figure 7. Comparing the average predicted cumulative GPA at different levels of LRU and building use (1 time per month vs. 4 times per month) under conditions of received CRI vs. not, and international student vs. U.S. resident. There is little to no difference between international students and U.S. residents compared at different levels of LRU and instruction.*
In the left column of Figure 8, the predicted difference in GPA between those with an LRU of 4 and those with an LRU of 1 varied by discipline. Veterinary Medicine students reached the undergraduate difference of 0.1. These students had a wider range of GPA than other graduate students. There was more variation for undergraduates, but those without CRI had a greater positive difference in all disciplines, ranging from about 0.05 to 0.1. The negative differences for those who did report CRI were in Mathematics & Computer Science and Business & Economics, both negative with about 75% probability. These two disciplines also had a smaller positive difference than other disciplines if they did receive CRI (0.05).

Because exploratory plots showed a distinct grouping with zero usage, and because the average of the GPAs of that group seemed to be higher than those with a small amount of usage, a dummy variable for zero usage was included in the model, which could vary by discipline. For this reason, the lower level of LRU in the comparison was 1 rather than zero usage.

These differences adjusted for gender (females average higher GPA); international status (difference varied by discipline, but was mostly neutral to negative); class level for undergraduates and whether a graduate student was in a no thesis, thesis, or dissertation program; the sampling strata; building use; and average cumulative GPA for class level and college.

For building usage, the average predicted difference in GPA (at 4 times per month compared to 1 time per month) was negative for both undergraduates and graduate students. This difference was the same for graduate international students and U.S. residents, but the difference was more pronounced between undergraduate international and U.S. residents (third line of Figure 7).
Figure 8. Average difference in predicted cumulative GPA, comparing an LRU of 1 time per month to 4 times per month, under conditions of CRI and no CRI (left), and comparing CRI to no CRI, while averaging over range of LRU. Thick segments show the middle 50% of plausible values, while thin segments show the middle 90%. If the entire 50% interval is greater than zero, then there is at least 75% probability that the difference is positive. If the entire 90% interval is greater than zero, then there is at least 95% probability that the difference is positive.
A simplified model was used for the scatter plots in Figure 9, leaving out CRI, building usage, residency, gender, and class level, and using only the zero use indicator and log2 of LRU as predictors, with varying slopes for both in the discipline/student type groups, and including the sampling strata (see Appendix for details).

The points in Figure 9 indicate the combination of the cumulative GPA (y-axis) and the LRU (on a logarithmic scale x-axis) for each individual respondent. The lines show how the GPA on average increased as the LRU doubled, with the slope of the line varying by discipline. The gray band indicates the 95% credible interval for the estimate of the slope of the line. The wide scatter of the points above and below the line at every level of LRU demonstrate that library resource usage did not explain very much of cumulative GPA and that there was a rather weak correlation. The angle of the line upward at zero, which varied by discipline, showed that there were students who reported no library resource usage yet still achieved high grades, more so in undergraduate disciplines, but also in the graduate disciplines of Business & Economics, Mathematics & Computer Science, and Physical Sciences & Engineering. The slope of the line tended to be steeper for most undergraduate disciplines and Vet Med students as they had a wider range of GPA.
Figure 9. Scatter plot of cumulative GPA vs. LRU on a logarithmic x-axis, with points representing each respondent’s combination. Lines are from the predictions from the model, drawn over the range of data for each discipline. Different intercepts for the sampling strata result in some jiggle in the lines. A zero-use indicator results in a higher prediction of GPA for students who report zero library resource usage and a steeper slope for students who do use library resources. Gray bands represent uncertainty in the estimate of the slope (95% credible intervals).
Discussion

Librarians would not expect the correlation between students’ success and their use of the library to be uniform across disciplines; multilevel modeling provides a way to estimate the varying strength of those correlations and levels. Rather than trying to reject a one-size-fits-all null hypothesis, this analysis was descriptive, providing estimates along with the uncertainty of those estimates.

The descriptive approach allows us to recognize the possibility that some students don’t use the library and yet do just fine in terms of their GPA. Putting this aside, when students do use library resources, the cumulative GPA is higher on average as the LRU index doubles, although this is not so for building usage. For undergraduates, this correlation between LRU and GPA is true only if they did not report receiving CRI in the last year. How can we make sense of such results?

As much as we would like to believe that use of library resources leads to higher GPAs, the widely scattered points in the scatterplots (Figure 9) in this study make clear that LRU does not explain very much of GPA. The results are very similar to other analyses such as those from Soria, Fransen, and Nackerud, which show a small association between use of library resources and GPA (an increase of .02 in $R^2$). This is not surprising. As Meehl (1990) put it, the “crud factor” means that in the social sciences “everything correlates to some extent with everything else.”

Disciplinary Differences: As expected, results varied by discipline. In all disciplines, if students believed the library contributed to their success (LCS), they tended to use resources (LRU) more. In some disciplines, this tendency was more pronounced; for example, LCS predicted more of a difference in LRU for graduate students in Veterinary Medicine (200% higher LRU) and Agriculture & Life Sciences, and less difference for graduate students in other subject areas (as low as 75% higher LRU). Results for undergraduate students also varied by discipline, again with LCS predicting more of a difference in LRU in some subject areas. In addition, disciplinary results differed by student type, as subject areas with the highest or lowest difference were not the same for undergraduate and graduate students. We might posit what these results might mean but that is beyond the scope of what the survey can tell us. Finally, predictions of CRI’s effect for LRU did not vary much across disciplines.

The predicted differences in GPA with higher levels of LRU did not vary much across disciplines if there was no CRI; for undergraduates, the difference was 0.05 in two of the disciplines, compared to 0.1 for all other disciplines. If CRI was received, the differences varied more, ranging from -0.05 to 0.05, and were estimated with more uncertainty (wider intervals) than the no CRI differences.

Student Attitudes on Success: The survey respondents who said that LRU contributed “very much” to their academic success [LCS] tended to be in the higher end of both LRU and GPA. This study found that frequency of library building usage is predictive of undergraduates believing that the library contributes “very much” to their success, more than LRU, yet data results indicate building use is not predictive of higher GPA. How do students define success and why do undergraduates think that library building use contributes to it? Similarly, why do results from graduate students disagree with these findings?

Educational research has found compelling linkages between students’ sense of belonging with their retention and success (Strayhorn, 2019, and summarized in Oliveira, 2017) and this is particularly so with students of color and other minority communities. For undergraduates, might it be that repeated library building use leads to their greater familiarity and comfort, contributing to a sense of “belonging” at the library where one can “fit in” and engage in various scholarly behaviors and activities likely to lead
to success? Our study found that international students were 6-8% more likely than other student groups to agree that the library’s resources and services contribute “very much” to their success, and they also reported higher LRU.

**Instruction (CRI), Library contributed to success (LCS) & Library resource use (LRU)**

Many library value studies have sought to examine the impact of library instruction on student GPA and other success measures. Our study defined library instruction as course-related instruction (CRI), or as the survey put it, a librarian coming to one of their classes. The results of this study are similar to others in that they show modest, if hard to explain, positive correlations after adjusting for LRU between library instruction and student success, defined as student beliefs in whether they felt the library contributed very much to their success (LCS), and negative correlations with GPA, again after adjusting for LRU.

In other words, students who had attended a CRI session were more likely to have a favorable attitude about the library’s role in their success, even though their GPA may not have been similarly impacted.

Are classes that receive instruction more likely to be harder or easier to get a good grade in? Although most disciplines showed a slightly negative or neutral result, if there was a positive effect of library instruction on GPA, it may be diluted by the effect of all other college experiences the student receives.

Receiving CRI did strongly predict higher LRU for many student groups, especially international students. These results may be explained by a strong likelihood for courses receiving CRI to have assignments or expectations that require LRU for student assignments or papers. This may also suggest students learn through CRI about relevant resources as well as the databases and strategies to explore for finding additional resources on their own.

While these results are intriguing and open to interpretation, are these types of library value studies emphasizing big-picture quantitative data the most productive ways of looking at the value of library instruction? For example, it is difficult for quantitative studies to get at the quality issues that matter in learning and teaching, such as the quality of the collaboration between librarian and course instructor; the type of learning strategies employed; the level of instruction offered and how well it aligns with student and course needs; the expertise of the librarian as teacher; and so on.

**Recommendations**

It must be recognized that cumulative GPA may be a blunt instrument to measure the effect of library usage or instruction on student success. In what proportion of classes is it plausible that library usage or instruction would help to increase the student’s grade? The effect could be a drop in the pond in a cumulative GPA.

To more fully understand and measure the relationship between library building and resource use and student success, it would be helpful to examine how students define success and how the library helps them. A longitudinal study such as a diary study or panel survey could gather such qualitative information, while allowing students to be their own controls, using multilevel models with multiple observations nested in each student.

Whitmire’s study of library usage is an excellent example of a much fuller data set than ours. It has a range of variables, including background characteristics, college experiences, and library activities, and
importantly, it covered three years with the same students. A multilevel model could combine all the
data with the three years’ observations nested within each student and students nested within
institutions.

Normand47 recommends increasing the number of observations per person over increasing the number
of people in the study. With students as their own control, all aspects of their background, personality,
social/job life, study habits, and abilities can be controlled for. Using outcomes linked more tightly to the
treatments and focusing on variations over time in individual performance rather than averages, such a
study could provide much richer information about how and why the library can help student success.

**Conclusion:** Student attitudes and beliefs about their own library use matter, whether or not their use
can objectively be shown to have been positively linked to success outcomes such as higher GPA. This is
particularly so with undergraduates in our study. GPA may not be the best measure of whether the
library is an aid to student success, although there is a correlation between LRU and GPA. Higher LRU,
building use, and CRI does predict student attribution of academic success to the library; by self report,
it helps some students. It may be of more help to some students than to others either because it is
discipline dependent or because availability of alternatives varies. It’s important also to recognize that
students are diverse and do not represent a homogeneous block, as seen with the differences between
US residents and international students in this study. Including more student-centered measures of
specific ways the library can aid the student, including student-defined measures of success, and
including more observations per student might bring the issue of library impact on student success into
sharper focus.

---

1 Connaway, Lynn Silipigni, William Harvey, Vanessa Kitzie, and Stephanie Mikitish. "Academic library impact:


5 Ibid., p. 14.

6 Wong, Shun Han Rebekah, and T. D. Webb. 2011. “Uncovering Meaningful Correlation between Student
70. [https://doi.org/10.5860/crl-129](https://doi.org/10.5860/crl-129).

7 Soria, Krista M., Jan Fransen, and Shane Nackerud. 2013. “Library Use and Undergraduate Student Outcomes:


12 Soria, Fransen, and Nackerud. “Library Use and Undergraduate Student Outcomes.”


18 Wong and Cmor. “Measuring Association between Library Instruction and Graduation GPA,” 147–64.


22 Wong and Cmor. “Measuring Association between Library Instruction and Graduation GPA,” 147–64.


25 For the graduate student respondents, the cumulative GPA did not include Spring 2016, while Spring 2016 was included for the undergraduates.
Our study asked students themselves to indicate whether they received library instruction. While this approach is not problem-free, as it does not delve into quality or level of instruction and relies as well on student memory, it avoids the problems of inferred CRI attendance.
