acrl 2021
virtual

AN HOUR OF YOUR TIME FOR YEARS OF SUCCESS?

A Longitudinal Analysis of How Library Curriculum–Integrated Instruction May Mitigate Disparities in Undergraduate Academic Performance

Felichism Kabo, Doreen Bradley, and Laurie Alexander*

This study examines the impact of library curriculum–integrated instruction (CII) on short–term and long–term academic performance for undergraduates. Generally, library CII is considered essential to the academic performance and success of students in higher education. We propose that library CII may also be used to mitigate the factors that are associated with disparities in academic performance among students. Specifically, we evaluate the mediating role that library CII has on potential digital disparities associated with online access to library–licensed content. We employ large–scale library usage and institutional student data and apply longitudinal statistical modeling in our analysis. Specifically, we run panel linear mixed effects regression models of the association between library CII and student performance, and the study results confirm a positive and statistically relationship between library CII and short– (semester) and long–term (cumulative) student grade point average (GPA). Further, for long–term academic performance, participation in a course that has library CII has a similar effect size or magnitude of impact as digital access to library–licensed content, our proxy for potential digital disparities among students. The study results affirm the value and importance of library CII for student performance and invite us to engage in conversations on new approaches to library CII that are informed by multiple perspectives such as those of students, faculty, administrators, and librarians.

* Felichism Kabo, Assistant Research Scientist, Institute for Social Research, University of Michigan, email: fkabo@umich.edu; Doreen Bradley, Director of Learning Programs and Initiatives, University of Michigan Library, email: dbradley@umich.edu; Laurie Alexander, Associate University Librarian for Learning and Teaching, University of Michigan Library, email: lauriea@umich.edu

INTRODUCTION

Information literacy skills are essential to the academic performance and success of students in higher education (McGuinness 2011b). Library curriculum–integrated instruction (CII) is the embedding of information literacy in academic courses (McGuinness 2011b). CII typically consists of workshops or training sessions that are designed to teach library research and technology skills to students to enable them to complete their academic work. These workshops are tailored to meet the specific learning outcomes of courses and programs, most often focusing on requirements of assignments and projects. "Curriculum–integrated" refers to the mode of delivery of these workshops or training sessions, where they are integrated into subject–specific teaching (e.g., introductory physics course) based on the assumption that this will help students acquire the targeted skills more effectively than when they are taught as a separate course (VanScoy and Oakleaf 2008, McGuinness 2011a). Bowles–Terry (2012) concludes that students who participate in library instruction have higher GPAs, and that library instruction is particularly effective when students participate in multiple years during their undergraduate careers.

Library instruction potentially benefits students in the short–term by showing them efficient information location and retrieval, and in the long–term by teaching them the logics of information access and critical thinking, hence lifelong learning skills. Library instruction could also mitigate student performance disparities. We present findings from an IMLS–funded study of student performance done by merging large–scale institutional and library usage data. Therefore, this study is aligned with the ACRL's call for quantification of the relationship between library usage and student success (Lynn Silipigni, William, and Vanessa 2017). Most studies that have examined the relationship between library usage and undergraduate student success have been survey–based and are characterized by small samples and low response rates. For example, a recent study targeting 2,952 undergraduates had a response rate that was less than 17% (Anderson and Vega García 2020). Another study attempted to examine the links between increased exposure to information literacy training and students' perceived competency and confidence at two schools of nursing, but was only able to report the results from one site due to very low response rates at the other site (Farrell, Goosney, and Hutchens 2013). Finally, a survey distributed to 8,522 students as part of a study of library usage among transfer students had a very low response rate of 5.39% (Richter–Weikum and Seeber 2018). Low response rates and small samples affect the generalizability of study findings to broader groups of students at the study sites and elsewhere and make it more likely that the studies are not reproducible in other institutional settings beyond where the studies were performed. The small sample sizes in these studies also mean that the findings have low statistical power and are thus of questionable utility in making inferences about racial/ethnic minorities and other under–represented groups. Our study avoids these issues by using large–scale, longitudinal administrative data (library usage and institutional student data) that have sufficient statistical power for us to perform analyses for subgroups such as racial/ethnic minorities, and to examine differences across factors such as academic units.

The nationwide work–from–home mandates due to the COVID–19 pandemic have likely exacerbated the effects of the digital divide or disparities in internet and technology access. Examination of online access to library–licensed content may illuminate digital inequalities in access to reliable, high–speed internet, and technical expertise e.g., setting up and using a virtual private network (VPN). Learning analytics (LA) entails the measurement, collection, and analysis of student data and their contexts to better understand how to improve the learning process. Library usage data are a key element of the student learning process that are missing entirely from most LA studies. We propose that library learning analytics (LLA), the collection and analysis of library usage data and connecting these data with other institutional data on students to better understand how interacting with the library shapes the learning process and student outcomes, is a potentially powerful change agent. LLA has enormous potential to better articulate the value and importance of the library for student performance and can provide direction for how library services may be used to mitigate disparities in student outcomes. In this paper we present the results of an LLA study that examines the associations between library CII and student GPA in the short– (semester) and long–term (cumulative) for undergraduates enrolled at a public, research institution over a six–term period. Further, we compare the magnitude of the impact of CII on student performance relative to online access to library–licensed content, and hence potential digital disparities.

METHODS

We examine the impacts of library CII on performance, and its potential mitigation of disparities in digital access to library–licensed content that is correlated with student performance. We use a large unstructured library EZproxy use log comprising over 800 million records going back to 2016, and over 5 million records of longitudinal institutional data on student demographics and course performance for the period from fall 2016 through winter 2019. Using Python scripts and regular expressions, the library log is structured, cleaned, and stored in a secure server that has strict controls on data egress and ingress. The structured library data are merged with the institutional student data using Structured Query Language (SQL) scripts, and the merged records are stored in a relational database on the secure server. The sample is all undergraduates (N = 45,273) enrolled at the University of Michigan (U–M) from fall 2016 through winter 2019. Our analysis focused on students who were enrolled in at least five semesters over the six–semester study period. Limiting the analysis to this group of students enabled us to have more balanced panels for the longitudinal regressions. With balanced data, we can observe the same unit in most or all time periods in the study which reduces the noise introduced by individual–level heterogeneity. The models resulting from this analysis are less likely to be unduly influenced by outliers such as a student who drops out of the university after one term of enrollment.

Variables

The outcomes or dependent variables are student grade point average (GPA) in the short– (semester) and long–term (cumulative). Semester GPA ("SEM_GPA") and cumulative GPA ("CUM_GPA") are both continuous variables coded on 0–4.4 and 0–4.314 scales, respectively. The predictor or independent variable ("Participated in Library CII") is a dichotomous variable that is coded 1 if a student took a course that had library CII in an academic term and is coded 0 otherwise. At U–M, library instructors typically teach 800–900 CII sessions per year–most of which are one–time sessions ranging from 50 minutes to 1.5 hours–to approximately 15,000 undergraduate and graduate students.

The covariates or control variables are library access, socio–demographic, and academic variables. The variables are operationalized as follows:

- "**Ever EZproxy session in term**" is a dichotomous variable coded 1 if a student had at least one EZproxy session in an academic term and coded 0 otherwise. This captures online access to library–licensed content and is our proxy for potential digital disparities among students.
- Sex is captured by the dichotomous variable "**Sex**" (1 = Female, 2 = Male).
- High school GPA ("**HS_GPA**") is a continuous variable on a 0–4 scale.
- Race/ethnicity is captured by the categorical variable "**RACE/ETHNICITY**" (1 = White, 2 = Asian, 3 = Black, 4 = Hispanic, 5 = Two or More, 6 = Other, 7 = Not Indicated)
- The indicator for whether a student has first generation status is the categorical variable "**FIRST GEN STATUS**" (1 = First Gen, 2 = Not First Gen, 3 = Don't Know)
- The student's family income is captured by the categorical variable "**FAMILY INCOME**" (1 = More than $100,000, 2 = Less than $25,000, 3 = $25,000–$49,999, 4 = $50,000–$74,999, 5 = $75,000–$99,999, 6 = Don't Know, 7 = Missing)
- The categorical variable "**PARENTAL EDUCATION**" (1 = Associate's degree, 2 = Elementary School only, 3 = Less than High School, 4 = High School diploma, 5 = Some College, 6 = Bachelor's degree, 7 = Master's degree, 8 = Professional Doctorate, 9 = Doctorate, 10 = Post Doctorate, 11 = Don't Know) captures the educational attainment of the student's parents
- The student's class level is indicated by the categorical variable "**CLASS LEVEL**" (1 = Freshman, 2 = Sophomore, 3 = Junior, 4 = Senior).

Statistical Analysis

Students are rarely treated as independent with respect to their academic performance and success. Academic outcomes are a function of factors that are common to all students (e.g., gender, race, etc.) that are treated as

"fixed effects" and factors that cluster student behaviors and outcomes (e.g., academic units and disciplines, etc.) that are regarded as "random effects." At the student–level, other random effects include unobserved, time invariant factors such as motivation, grit, and persistence. Therefore, we run longitudinal or panel linear mixed effects (LME) regression models that account for both fixed and random effects. Regression modeling and analysis was done using the statistical software Stata MP release 16 (StataCorp 2019). Post–estimation likelihood–ratio tests comparing each LME model with a one–level ordinary linear regression were significant, lending support to the decision to apply the LME models for the regression analysis.

RESULTS

Descriptive Statistics

Over the study period the percentage of students who took a course that had library CII in each academic term was between 21–26% (Table 1).

Focusing on the latest academic term in the study period (winter 2019, we see that there is more variation by socio–demographic and academic factors (Table 2) than there is between semesters (Table 1). For example, more females (25%) than males (16%) took a course that had library CII in winter 2019. There is a notable decrease in the percentage of students taking courses with library CII as they advance through the class levels from freshman (36%) to senior (14%). However, this is to be expected as more information literacy instruction takes place during individual consultations with librarians for upper–level undergraduates as their research interests become more specialized.

TABLE 1

Percentage of students who took course with library CII by semester, Fall 2016–Winter 2019

Academic Term	Number of Students	Library CII	Percent Library CII
FA 2016	28,682	6,973	24%
WN 2017	27,408	5,798	21%
FA 2017	29,161	7,566	26%
WN 2018	27,852	6,227	22%
FA 2018	29,743	6,498	22%
WN 2019	28,371	5,819	21%

TABLE 2

Student participation in courses with library CII by demographic and academic factors, Winter 2019

Variable	Category	Number of Students	Library CII	% Library CII
First Gen Status	Don't Know	47	17	36%
	First Gen	3,893	953	24%
	Not First Gen	24,431	4,849	20%
Family Income	Less than $25,000	1,507	372	25%
	$25,000–$49,999	2,215	499	23%
	$50,000–$74,999	2,011	420	21%
	$75,000–$99,999	2,075	467	23%
	More than $100,000	13,959	2,765	20%
	Don't Know	515	93	18%
	Missing Income Information	6,089	1,203	20%
Class Level	Freshman	2,560	931	36%
	Sophomore	6,408	1,805	28%
	Junior	7,133	1,332	19%
	Senior	12,270	1,751	14%

TABLE 2

Student participation in courses with library CII by demographic and academic factors, Winter 2019

Variable	Category	Number of Students	Library CII	% Library CII
Race	Asian	5,832	1,099	19%
	Black	1,269	319	25%
	Hispanic	1,900	451	24%
	White	16,613	3,339	20%
	2 or More	1,303	302	23%
	Other	46	12	26%
	Not Indic	1,408	297	21%
Sex	Female	14,211	3,518	25%
	Male	14,160	2,301	16%
Academic Unit	Undergrad Music, Thtre & Dance	717	65	9%
	Undergraduate Architecture	181	38	21%
	Undergraduate Art and Design	524	224	43%
	Undergraduate Business Admin	1,801	174	10%
	Undergraduate Dental Hygiene	101	3	3%
	Undergraduate Education	126	12	10%
	Undergraduate Engineering	6,314	780	12%
	Undergraduate Information	260	15	6%
	Undergraduate Joined Deg Prog	10	2	20%
	Undergraduate Kinesiology	954	243	25%
	Undergraduate L S & A*	16,422	3,755	23%
	Undergraduate Nursing	607	453	75%
	Undergraduate Pharmacy	55	2	4%
	Undergraduate Public Health	157	37	24%
	Undergraduate Public Policy	142	16	11%
Parental Education	Elementary School only	189	42	22%
	Less than High School	292	62	21%
	High School diploma	1,674	402	24%
	Some College	806	175	22%
	Associate's degree	844	214	25%
	Bachelor's degree	7140	1418	20%
	Master's degree	9001	1674	19%
	Professional Doctorate	5,082	997	20%
	Doctorate	2096	392	19%
	Post Doctorate	—	—	—
	Don't Know	45	15	33%

**Literature, Science, and the Arts*

Blacks (25%) and Hispanics (24%) are more likely to take courses with library CII than Whites (20%), which likely reflects the library's success in establishing instructional relationships with area studies and student success programs. Finally, we note that there are large differences in the percentage of students taking courses with library CII across academic units ranging from a low of 3% (Dental Hygiene) to a high of 75% (Nursing).

Regression Analysis

The results show that library CII is positively correlated with short– (model 2) and long–term (model 4) student GPA, and these associations are statistically significant. For both semester and cumulative GPA, the addition of the library CII predictor variable (models 2 and 4) to the baseline models that do not have the predictor (models 1 and 3) attenuates the effect of digitally accessing library–licensed content (having at least one EZproxy session in an academic term). Adding the library CII predictor to the model reduces the size of the coefficient (beta) for having at least one EZproxy term in an academic term by 31.4% and 17.6% for semester and cumulative GPA, respectively. Moreover, in the long–term, CII ($\beta = 0.014$, $p < .001$) has an impact of roughly similar magnitude to having at least one EZproxy in an academic term ($\beta = 0.015$, $p < .001$), suggesting that it could be used to mitigate the potential impact of digital factors that contribute to disparities in student performance.

TABLE 3

Panel LME regressions on semester and cumulative GPA, FA 2016–WN 2019 (min. 5 semesters)

	(1)	(2)	(3)	(4)
VARIABLES	**SEM_GPA**	**SEM_GPA**	**CUM_GPA**	**CUM_GPA**
Took course with Library CII		0.0712***		0.0136***
		(0.00494)		(0.00136)
Ever EZproxy session in term	0.141***	0.0967***	0.0182***	0.0150***
	(0.00433)	(0.00401)	(0.00113)	(0.00112)
Sex (1=Female, 2=Male)	–0.0851***	–0.0790***	–0.0624***	–0.0613***
	(0.00812)	(0.00766)	(0.00667)	(0.00664)
HS_GPA	0.0459***	0.0416***	0.0364***	0.0361***
	(0.00358)	(0.00338)	(0.00297)	(0.00296)
RACE/ETHNICITY (reference = White)				
Asian	0.0175+	0.0167+	0.0315***	0.0313***
	(0.0104)	(0.00985)	(0.00860)	(0.00857)
Black	–0.372***	–0.371***	–0.337***	–0.336***
	(0.0190)	(0.0180)	(0.0157)	(0.0156)
Hispanic	–0.169***	–0.169***	–0.167***	–0.167***
	(0.0170)	(0.0160)	(0.0140)	(0.0140)
2 or More	–0.106***	–0.0995***	–0.0624***	–0.0622***
	(0.0192)	(0.0181)	(0.0159)	(0.0159)
Other	–0.308**	–0.300**	–0.232**	–0.234**
	(0.0970)	(0.0916)	(0.0812)	(0.0809)
Not Indicated	–0.0483*	–0.0345+	–0.0371*	–0.0347*
	(0.0205)	(0.0193)	(0.0169)	(0.0169)
FIRST GEN STATUS (reference = Not First Gen)				
First gen	–0.176***	–0.165***	–0.140***	–0.141***

TABLE 3

Panel LME regressions on semester and cumulative GPA, FA 2016–WN 2019 (min. 5 semesters)

	(1)	(2)	(3)	(4)
VARIABLES	**SEM_GPA**	**SEM_GPA**	**CUM_GPA**	**CUM_GPA**
	(0.0274)	(0.0259)	(0.0227)	(0.0227)
Don't Know	–0.197+	–0.207*	–0.166+	–0.160+
	(0.104)	(0.0978)	(0.0872)	(0.0869)
FAMILY INCOME (reference = More than $100,000)				
Less than $25,000	–0.117***	–0.112***	–0.0839***	–0.0831***
	(0.0199)	(0.0188)	(0.0164)	(0.0164)
$25,000–$49,999	–0.0862***	–0.0873***	–0.0702***	–0.0708***
	(0.0172)	(0.0162)	(0.0142)	(0.0142)
$50,000–$74,999	–0.0653***	–0.0606***	–0.0498***	–0.0496***
	(0.0164)	(0.0155)	(0.0135)	(0.0135)
$75,000–$99,999	–0.0503**	–0.0460**	–0.0381**	–0.0389**
	(0.0154)	(0.0145)	(0.0127)	(0.0127)
Don't Know	–0.0279	–0.0327	–0.000916	–0.00122
	(0.0281)	(0.0265)	(0.0229)	(0.0229)
Missing Income Information	0.000839	0.00132	0.00761	0.00742
	(0.00999)	(0.00942)	(0.00822)	(0.00819)
PARENTAL EDUC. (reference = Associate's degree)				
Elementary School only	–0.0680	–0.0808	–0.0940*	–0.0945*
	(0.0527)	(0.0497)	(0.0437)	(0.0435)
Less than High School	–0.00714	–0.00156	–0.0292	–0.0290
	(0.0460)	(0.0434)	(0.0379)	(0.0377)
High School diploma	0.0207	0.0146	0.00218	0.00322
	(0.0305)	(0.0288)	(0.0253)	(0.0252)
Some College	0.00684	–0.00861	–0.0299	–0.0280
	(0.0328)	(0.0309)	(0.0272)	(0.0271)
Bachelor's degree	–0.0872***	–0.0853***	–0.0819***	–0.0815***
	(0.0118)	(0.0111)	(0.00968)	(0.00965)
Master's degree	–0.0445***	–0.0438***	–0.0429***	–0.0424***
	(0.0110)	(0.0104)	(0.00907)	(0.00904)
Professional Doctorate	–	–	–	–
Doctorate	–0.000333	–0.000905	0.00946	0.00943
	(0.0169)	(0.0159)	(0.0140)	(0.0139)
Post Doctorate	–	–	–	–
Don't Know	–	–	–	–

TABLE 3

Panel LME regressions on semester and cumulative GPA, FA 2016–WN 2019 (min. 5 semesters)

	(1)	(2)	(3)	(4)
VARIABLES	**SEM_GPA**	**SEM_GPA**	**CUM_GPA**	**CUM_GPA**
CLASS LEVEL (reference = Freshman)				
Sophomore	–0.0321***	–0.0131+	–0.000872	8.34e–05
	(0.00747)	(0.00683)	(0.00193)	(0.00189)
Junior	–0.0425***	–0.0100	0.00222	0.00423*
	(0.00752)	(0.00694)	(0.00197)	(0.00194)
Senior	–0.0207**	0.0281***	0.0206***	0.0233***
	(0.00791)	(0.00734)	(0.00210)	(0.00208)
Constant	3.500***	3.498***	3.518***	3.514***
	(0.0405)	(0.0396)	(0.0317)	(0.0317)
Observations	79,487	78,967	79,487	78,967
Number of groups (schools/ colleges)	15	15	15	15
Standard errors in parentheses				
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1				

DISCUSSION

Results from the study support previous findings on positive associations between library CII and student performance. The results also support our proposition of the potential for library CII to mitigate potential disparities in academic performance accruing from digital factors such as online access to library–licensed content. Correlations between CII and improved student performance during a given semester could serve as an impetus to integrate CII more fully into curricula in our degree programs to enhance information literacy skills throughout students' academic careers. Soria et al (2013) found increased library use correlates with increased measures of student success. As CII often opens the door to libraries, particularly for students who did not use libraries prior to college, exploring new methods to scaffold library instruction into our schools and colleges could be particularly valuable.

We find it noteworthy that library CII has significant impacts on short– and long–term student outcomes even though most CII sessions are about 1–1.5 hours long, which means that they have a relatively brief duration in the context of the length of an academic term. In other words, library CII has a major multiplier effect on student performance. The large–scale data, statistical modeling, and findings of this study suggest that academic libraries can engage in rigorous and thoughtful LLA activities using frameworks that protect student privacy and confidentiality and in so doing make valuable contributions to existing LA ecosystems at their parent institutions (Alexander, Bradley, and Varnum 2018).

FUTURE RESEARCH

The study findings suggest multiple potential lines of future inquiry into the links between library CII and student performance. For example, areas for future research could include analyses of students who have two or more CII sessions during their undergraduate coursework, or CII in a course within their major, and how these translate to higher GPAs at graduation, and other measures of student success.

REFERENCES

Alexander, Laurie, Doreen R. Bradley, and Kenneth J. Varnum. 2018. "On the Road to Learning Analytics: The University of Michigan Library's Experience with Privacy and Library Data." In *Using Digital Analytics for Smart Assessment*, edited by Tabatha Farney, 83–93. Chicago, IL: American Library Association.

Anderson, Linda L., and Susan A. Vega García. 2020. "Library Usage, Instruction, and Student Success across Disciplines: A Multilevel Model Approach." *College & Research Libraries* 81 (3):459–491. https://doi.org/10.5860/crl.81.3.326.

Bowles–Terry, Melissa. 2012. "Library Instruction and Academic Success: A Mixed–Methods Assessment of a Library Instruction Program." *Evidence Based Library and Information Practice* 7 (1):82–95. https://doi.org/10.18438/B8PS4D.

Farrell, Alison, Janet Goosney, and Karen Hutchens. 2013. "Evaluation of the Effectiveness of Course Integrated Library Instruction in an Undergraduate Nursing Program." *Journal of the Canadian Health Libraries Association / Journal de l'Association des bibliothèques de la santé du Canada* 34 (3):164–175. https://doi.org/10.5596/c13–061.

Lynn Silipigni, Connaway, Harvey William, and Kitzie Vanessa. 2017. *Academic Library Impact : Improving Practice and Essential Areas to Research*. Chicago, IL, UNITED STATES: Association of College and Research Libraries.

McGuinness, Claire. 2011a. "The evolving role of the teaching librarian." In *Becoming Confident Teachers*, edited by Claire McGuinness, 1–31. Chandos Publishing.

McGuinness, Claire. 2011b. "Introduction." In *Becoming Confident Teachers*, edited by Claire McGuinness, xv–xviii. Chandos Publishing.

Richter–Weikum, Erin, and Kevin Seeber. 2018. "Library experiences of transfer students at an urban campus." *Student Success* 9 (2):63–71. http://dx.doi.org/10.5204/ssj.v9i2.411.

Soria, Krista M., Jan Fransen, and Shane Nackerud. 2013. "Library Use and Undergraduate Student Outcomes: New Evidence for Students' Retention and Academic Success." *portal: Libraries and the Academy* 13 (2):147–164. https://doi.org/10.1353/pla.2013.0010.

Stata Statistical Software: Release 16. StataCorp LLC, College Station, TX.

VanScoy, Amy, and Megan J. Oakleaf. 2008. "Evidence vs. Anecdote: Using Syllabi to Plan Curriculum–Integrated Information Literacy Instruction." *College & Research Libraries* 69 (6). https://doi.org/10.5860/crl.69.6.566.