

School Library Research in the Real World—What Does it Really Take?

Joette Stefl-Mabry, Ph.D.

University at Albany, State University of New York

Michael S. Radlick, Ph.D.

Learning Technology Visions, LLC

ABSTRACT

School libraries are perceived to have a significant effect on student achievement. The reality is that evidence supporting the effects of school libraries on student achievement remains unconvincing to many serious researchers. In this paper, we provide a systematic review of 25 years of school library research examining student achievement. Results indicate that of over 260 studies, fewer than 27 approach the minimum requirements of research design. The unembellished truth is that most school library studies suffer from limitations of design, measurement, and analysis. To address such limitations, we built multiple statistical models based on six years of school-level data reflecting all public schools in New York State. We highlight key challenges of quantitative research: design, indicators, measurement and analysis approaches as they apply to ours and other school library research and share initial results from our study examining the causal relationships among school librarians, resources, activities and student achievement.

Keywords: School Libraries, Student Academic Achievement, Causality, Causal Studies, Research Design

INTRODUCTION

Since the late 1990s researchers have written that school libraries, if properly staffed and adequately resourced, have a positive impact on student achievement. By 2001 (Lance) researchers persuasively proclaimed that “the evidence” was “mounting!”. Lance reported that there were “about 75 studies on the impact of school library media programs on academic achievement” and that early studies had demonstrated “the value of the mere presence of a professionally trained and credentialed library media specialist” (Lance, 2001, p. 1). Although earlier Lance (2001) had cautioned that such correlations “beg the question of what the school librarians are doing that makes a difference” he explained that more recent studies had focused on school librarians as “creators and collaborators” and revealed that: “students perform better academically where the library media specialist:

- Is part of a planning and teaching team with the classroom.
- Teaches information literacy.
- Provides one-to-one tutoring for students in need” (p. 2).

Since then, more studies, based largely on correlation, assert that school libraries improve academic outcomes for children (Lance, Rodney, & Hamilton-Pennell, 2003), that “Powerful Libraries Make Powerful Learners” (Lance, Rodney, & Hamilton-Pennell, 2005), that students, teachers and administrators benefit from school libraries (Lance, Rodney, & Russell, 2007), and that investments in school libraries “pays off” in terms of student achievement (Lance & Schwarz, 2012).

When such studies have been criticized for attempting to establish a cause-and-effect relationship between characteristics of library media programs and achievement (Lance, 1994), Lance (2001) explains that “a stronger claim that correlations reflect cause-and-effect” had been established in the Alaska, Pennsylvania, and Colorado studies by including data on schools “i.e., per pupil spending, teacher-pupil ratio, various teacher characteristics” and their communities “i.e., poverty levels, racial/ethnic demography, adult educational attainment” and that the additional variables addressed “most, if not all, of the stronger arguments that could otherwise be made to discount the consistent

findings of this line of research” (p. 5). However, as anyone who has ever taken a course in statistics will confirm, correlation does not equal causation. There is no additive value whereby multiple correlations equal causation.

To date there have been 21 state-level studies that “confirm that school libraries support student achievement” (see infographic at http://www.ala.org/aasl/sites/ala.org.aasl/files/content/aaslissues/advocacy/AASL_infographic.pdf). The unquestioning acceptance of these “state studies” – none of which has been peer reviewed (although summaries of the findings have been reported in peer reviewed journals)—has shaped the way the school library profession perceives school libraries as well as the methodological design of subsequent school library research. It is argued that these studies provide, in the aggregate, strong, congruent evidence of the efficacy of school libraries in effecting student learning. However, as we reveal in this paper, this body of work, frequently referenced to argue that school librarians and school libraries are critical to student learning, is not as persuasive as its advocates believe. The field seems to be in a perpetual state of confirmation bias, with practitioners continuing to interpret all prior work as evidence that libraries unquestionably impact student achievement. As these authors have detailed in earlier work (Stefl-Mabry, Radlick, Armbruster, & Keller, 2016) the majority of school library studies over the past few decades have not even been based on data.

If school library research is to advance an agenda of rigorous empirical scrutiny, then an honest evaluation of the quality of school library research that has been conducted so far is necessary. Unlike other areas of research where multiple studies have been successfully replicated results across many circumstances, the research focused on the linkage between school libraries and student achievement suffers from multiple problems and inconsistencies that result in weak evidence, at best, supporting school library effects on achievement. The purpose of this paper is to:

1. Provide a topological and methodological review of school libraries and student achievement research from 1990-2017.
2. Review key challenges of quantitative research: research design, indicators and measurement, and analysis approaches.
3. Recommend alternatives to address weaknesses in school library research design and analysis and specify a better understanding of the necessary conditions for causal inference in a quantitative study.
4. Highlight initial results from our three year 2015-18 study (IMLS Grant RE-04-15-0081-15) which examines the causal relationships between school librarians, school library resources, school library activities and student achievement in public school libraries throughout New York State.

SCHOOL LIBRARY AND STUDENT ACHIEVEMENT RESEARCH 1990 THROUGH 2017

In 2016 we began a retrospective search to locate published reports of research in the field of school libraries and student achievement spanning a twenty-five-year period (1990-2015) for an American Educational Research Association conference paper: *Breaking down Information silos: Sharing decades of school library research with educational researchers* (Steffl-Mabry et al., 2016). The results of that study made us aware of the methodological weaknesses endemic to the field of school library research. As both authors participated in the two Causality: School Libraries and Student Success summits, we were also aware that the American Association of School Librarians (AASL) was seeking to articulate "...a national research agenda to investigate causal phenomena in school library instruction, resources, and services" (2014, p. 4).

Comparable to our 2016 findings, Morris & Cahill (2017) recently reported that research methods utilized in school library research "use less higher-level statistical analysis than methods used in other fields of study in education" (p.17). Although Morris & Cahill focused their analysis on articles written in two research journals, both of which deal exclusively with school libraries: *School Library Research* and *School Libraries Worldwide*, they found no evidence of "experimental or quasi-experimental designs" in their analysis of articles published from 2007 –July 2015 in either publication. Although Morris & Cahill's findings confirm ours, we remained puzzled by the multitude of voices from within the school library research community that proclaimed that "school libraries work" (Scholastic Research & Results, 2008), that "increasing numbers of investigations and improved methodology" bring "new credence and immediacy to this positive relationship" (Barrett, 2010, p. p. 136). and that "a wealth of research, or "big data" substantiates school libraries' impact on students' reading and writing" (Subramaniam, 2015, p. 17). These two conflicting and distinct narratives provided the impetus for us to conduct an intensive topical and methodological analysis of school library research and student achievement. Unlike Morris & Cahill (2017), who limited their search to two journals or Clyde (2004) who limited her review to library and information science publications, we extended our inquiry to include all fields and disciplines related to education and learning. Our goal was to provide an analysis of school library research to identify the:

- number of articles that have been published in peer reviewed scholarly journals focused on school library and student achievement.
- types of publications in which research involving school library and student achievement research are reported.
- range of topics that are targeted in school library and student achievement research.
- types of interventions that are examined in school library and student achievement research.
- methodological approaches that are used in school library and student achievement research studies.

METHODOLOGY

The literature review protocol for this study was based on the *What Works Clearinghouse Procedures and Standards Handbook*, Version 3.0 (What Works Clearinghouse, n.d.). Only research studies containing a "primary analysis of the effect of an intervention" were included in the first phase of the literature review rather than research theses or unpublished reports (p.7). Like Clyde (2004) we followed the belief that the main indexing and abstracting services in library and information science and education provide better coverage of published articles and papers, and that "the publication of findings is an important part of any research project" (p. 182).

Data Sources

The search was conducted in April 24, 2017 using the University at Albany's EBSCO interface. Scholarly research papers and conference papers published in English, were identified through a number of strategies. Online searches were undertaken of *Academic Search Complete*, *Education Source*,

Educational Administration Abstracts, ERIC and Library, Information Science & Technology Abstracts with Full Text. The search terms used for the database search are presented in Table 1 below:

Table 17: Search Terms Used for Database Search

“school libraries” OR “school librarians” OR “student achievement” OR “student academic achievement” OR “academic performance” OR “academic success” OR “scholastic achievement” OR “learning” OR “tests” OR “measures”	AND	“evidence” OR “empirical” OR “proof” OR “causal” OR “causality” OR “causation”	AND	Scholarly (Peer Reviewed Journals)
---	-----	--	-----	------------------------------------

Internet searches using public search engines also proved useful in phase two, to help us find research reports that had been cited in many of the peer-reviewed publications. There were instances, however, where the original report could not be located because the cited links were broken.

Search Results

Using the search terms listed in Table 1, a total of 230 scholarly research articles (published in English) were found. A breakdown of articles by journal discipline revealed that 85% (195) were published in library and information science journals and 14% (32) were published in educational journals. Only 1% (3) were published in journals other than library and information science or education, see Table 2 below.

Table 18: Publication by Discipline

Peer Review Journals	Number	Percent
LIS Publications	195	85%
Education Publications	32	14%
Other	3	1%
Total	230	100%

Topical Analysis of School Library and Student Achievement Research

To determine the focus area that had been investigated in each research article a topical analysis of the 230-peer reviewed publications was conducted. An adapted version of consensual qualitative research methodology (CQR) was used to examine the major topics represented (Anderson, Leahy, DelValle, Sherman, & Tansey, 2014; Hill, 2012; Hill et al., 2005). Codes were used to assign meaning to the descriptive and/or inferential information compiled by reading through each article multiple times (Miles & Huberman, 1994). Content analysis was performed to identify core concepts and consider the frequency with which they occurred. Codes were defined in terms of the focus of research of each publication.

Fifteen themes were identified as being central to the articles. Table 3 presents the topical areas of each of the 230 scholarly articles. Some articles covered multiple themes, thus total number of codes exceed the number of articles reviewed. See Table 3 below.

Table 19: Topic Areas Identified in School Library/Librarian Research and Student Achievement (230 publications)

Code	Code Count	Percentage
Advocacy	30	8%
Assessment/Evaluation	19	5%
Attitudes/Disposition	19	5%
Collaboration	40	10%
Curriculum/Instruction	58	15%

District Administration/Leadership	23	6%
Evidence Based Practice	13	3%
Library Programming	40	10%
Professional Development	6	2%
Professional Practice	29	7%
Research on Student Academic Achievement	33	8%
Research Summaries	34	9%
Resource Recommendations	18	5%
STEM/Science/Engineering/Math	7	2%
Technology	24	6%
Total Number of Code Counts*	393	100%

As Table 3 illustrates, 15% of the publications focus on curriculum and instruction and collaboration and professional development were the focus of 10% of the articles. Summaries of prior research were represented in 9% of the articles. Advocacy was the focus of 8% of the articles, and district administration/leadership and technology were the focus of 6% of the publications. Assessment, evaluation, attitudes and dispositions were the focus of 5% of the publications. Evidence based practice was the focus of 3% of the articles and STEM (Science, Technology and Engineering and Math) was the focus of 2% of the peer reviewed publications. Only 5% of the articles focused on school libraries and actual research measures of student achievement.

Final Selection of School Libraries and Student Achievement Research Publications

A structured matrix template was created by the authors to identify the major dimensions of each study to identify those studies which met the criteria for final selection. Only scholarly articles focused on PreK-12 education, containing a primary analysis of school library and student achievement data were included in the final selection. Matrix categories also included: purpose (study focus), population, methodology, data collection, research design, statistical analysis and results. The authors read through each paper and posted relevant information about the publication to a shared Excel file, consulting with each other if there were questions or issues concerning categorization. The matrix was refined as the publications were reviewed to ensure that information captured was relevant and accurate. The final categories used to review the publications are presented in Table 4. (A link to the final matrix can be provided by contacting the authors.)

Table 20: Matrix Categories

Matrix Categories
Year Published
Source
Author(s)
Level (National, State, Region, District)
Title of Study/Report
Years of Data Used
Purpose

Target and Unit of Analysis (District, School, Grade, Class or Student)
Population or Sample (Sampling Approach)
Group Characteristics (How Representative the Respondent(s)? Sample? Response Rate?)
Number of Cases
Contextual Factors Considered (Poverty, Demographics, etc.)
Data Collection Methodologies (Survey, Perceptual, Observation Data Systems)
Measures/Indicators Used
Design (Descriptive, Correlational, Causal, Experimental, Longitudinal, Meta-analysis)
Statistical Analysis Methodologies
Multiple Regression
ANCOVA
Effect Size
Causal Modeling
SEM (Structural Equation Modeling)
Path (Analysis/Coefficient)
Effects Found
APA Citation
Library & Information Science Publication
Education Publication
Peer Reviewed
Report
Dissertation

Rationale for Including Non-Scholarly Publications

As we discovered in our 2016 study although summaries of findings from state impact studies had been published in scholarly journals, the actual collection methods and explanation of how the data had been analyzed could only be found in the original unpublished reports (Stefl-Mabry et al., 2016). Thus, once again, we decided to include unpublished school library and achievement research reports, one widely cited dissertation and one book in the final collection totaling 266 publications, see Table 5 below.

Types of Publications

As Table 5 reveals, the vast majority of articles (73%) were published in library and information science journals. Only 12% of the articles were published in educational journals. Non-published reports made up 13% of the publications.

Table 21: Type and Total Number of Publications

Journal Type	Number	Percentage
Library & Information Science	195	73%

Educational Journals	32	12%
Other Journals	3	1%
Reports	34	13%
Book	1	0%
Dissertation	1	0%
Total	266	100%

Number of Publications in Library and Information Science Journals

A more detailed look at the number of articles published within library and information science journals reveals that 56% of the publications were published in practitioner-based journals. As Table 6 illustrates, 31% of publications were published in *Knowledge Quest* and 25% were published in *Teacher Librarian*. Both journals are designed primarily for school library practitioners. *School Library Research* (ISSN: 2165-1019), the scholarly refereed research journal of the American Association of School Librarians (AASL) and the successor to *School Library Media Research* (ISSN: 1523-4320) have published only 10% of school library and student achievement research papers. *School Libraries Worldwide*, the official professional peer-reviewed (refereed) research journal of the International Association of School Librarianship (IASL), contains only 9% of school library and student achievement research articles. Combined, these three “research” journals contain 19% of papers focused on school libraries and student achievement research.

Table 22: Library and Information Science Journals

Library & Information Science Journals	Number	Percentage
ACCESS [library source]	4	2%
Advances in Librarianship	1	1%
Annual Review of Information Science and Technology (ARIST)	1	1%
CILIP [Chartered Institute of Library & Information Professionals]	1	1%
Current Studies in Librarianship	1	1%
Emergency Librarian	3	2%
Evidence Based Library & Information Practice	4	2%
IFLA Journal	1	1%
Information Research	2	1%
Information Studies	1	1%
Investigación Bibliotecológica	1	1%
Journal of Academic Librarianship	1	1%
Journal of Access Services	1	1%
Journal of Education for Library & Information Science	2	1%
Journal of Librarianship & Information Science	1	1%
Journal of Librarianship & Information Services in Distance Education	1	1%
Journal of Library Administration	1	1%
Journal of the American Society for Information Science & Technology	1	1%
Knowledge Quest	60	31%
Library & Information Science Research	4	2%
Library & Information Update	2	1%

Library Issues	1	1%
Library Management	1	1%
Library Philosophy & Practice	2	1%
Library Quarterly	2	1%
Medium [Washington Library Media Association]	2	1%
New Library World	4	2%
Reference Librarian	1	1%
School Libraries in Canada	1	1%
School Libraries Worldwide	18	9%
School Library Media Research	15	8%
School Library Research	3	2%
Science & Technology Libraries	1	1%
Teacher Librarian	49	25%
Young Adult Library Services	1	1%
Total	195	100%

Number of Publications in Educational Journals

Published work related to school libraries and student achievement research is sparsely disseminated in educational research journals as is evident in Table 8 below.

Table 23: Education Journals

Education Journals	Number
AERA Conference Proceedings [American Educational Research Association]	3
AERA Open	1
Assessment in Education: Principles, Policy and Practice	1
Australian Journal of Education	2
California English	2
CATESOL Journal	1
Clearing House	1
De La Salle Lipa Integrated School Research Journal	1
Economics of Education Review	1
Education	1
Educational and Psychology Measurement	1
Educational Research Quarterly	1
Interactive Learning Environments	1
International Journal of Learning	1
International Schools Journal	1
Journal of Education and Practice	1
Journal of Information Technology Education	1
Language and Literacy Spectrum	1

Multicultural Review	1
New England Reading Association Journal	1
Radical Pedagogy	1
Reading Improvement	1
Research for Educational Reform	1
Research on Educational Technology	1
Review of Educational Research	1
Teaching Exceptional Children	1
TechTrends: Linking Research & Practice to Improve Learning	1
Zeitschrift für Pädagogik;	1
Total	32

Publications in Other Journals

A very small number of articles related to school libraries and student achievement were published in journals representing disciplines other than library and information science and education, see Table 8 below.

Table 24: Journals other than Library and Information Science and Education

Other Publications	Number
IFE Psychologia	1
Journal of Economic Behavior & Organization	1
Social Science Research	1
Total	3

METHODOLOGICAL ANALYSIS OF SCHOOL LIBRARY AND STUDENT ACHIEVEMENT RESEARCH

Rationale for Inclusion in Methodological Analysis

Studies met the inclusion criteria for methodological analysis if they were focused on PreK-12, peer reviewed or report, focused on school libraries/librarians, and examined achievement. The study also had to have actual data reflecting school library and student achievement variables that were measured or observed, and the study had to be the primary analysis of the data (not a compilation of studies, a summary or meta-analysis).

While many of the 266 publications included student achievement and/or student academic performance as key terms in its descriptor, studies were excluded from the review if the research:

- was contextualized in public or academic libraries or a higher education setting
- was a review or summation of prior research
- described examples of best practice (e.g., collaboration, programming, leadership, curriculum and/or instruction) with no empirical data
- described professional development initiatives
- was written in a language other than English.

Analysis of the 266 publications revealed that 30% (80) of the publications focused on school libraries (PreK-12) and its effect on student achievement while 70% (186) did not. Of the final list of 80 papers (44 scholarly peer reviewed publications 34 reports, one dissertation and one book), 55% were published in peer reviewed publications and 44% were not (see Table 9 below). Thirty-two (40%) of the papers were published in library and information science journals, 9 (11%) were published in educational

journals, and 3 were published in journals in disciplines other than library and information science and education, see Table 9.

Table 25: Papers Selected for Methodological Review

Publication Type	Number	Percentage
Library and Information Science Journal	32	40%
Educational Journal	9	11%
Report	34	43%
Dissertation	1	1%
Other Journals	3	4%
Book	1	1%
Total	80	100%

Methodological Analysis

Of the 80 studies, only 24 studies reflected a quasi-experimental research design and 3 used an experimental design with random assignment (RCT) in which attempts were made to assure comparability between experimental and control groups. These 24 studies used designs supported by techniques such as ANCOVA (n=3) or some form of regression (n=21). In addition, two of the studies that used regression also used structural equation modeling (path modeling). It is important to note that two of the RCT studies suffer from significant potential challenges to external validity due to subject assignment process and the small numbers of cases (in one study a total of 28 experimental and 26 control students) and in the other intact class groups from 14 experimental schools (n=272) and 10 control schools (n=411). Both studies did use individual student data, as did the third RCT study (Borkum, He, & Linden, 2012) that had a much larger sample with 191 treatment schools and 195 control schools, with a total study size of over 20,000 students. In contrast with the three RCT studies, among the quasi-experimental design studies, all three ANCOVA studies were at the state level, and 20 of the regression studies reflected large, state-level, or national populations, with attempts to control for student economic disadvantage and in some cases other variables. All of the quasi-experimental design studies were focused at the school, rather than the individual-student level.

Weaknesses of Prior Research

Examining the body of research focused on school libraries and student achievement as reflected in the 80 studies/reports, we identified the following specific weaknesses.

Lack of an Underlying Theory of Action

First, we could find no clearly articulated underlying theory of action focusing on school library effects on student achievement within a school. Researchers assume that there is a connection underlying any school library effect. However, the critical question left unanswered in the multitude of school library impact studies is “what is the theory of action for actual intervention of the school library and school librarian that is influencing student achievement?” Is it what the school library has (e.g. library resources such as the size of the collection, or type of resources such as books, eBooks, Chrome books, or other computers)? Is it what the school librarian does or does not do (e.g. teaching critical thinking skills, inquiry-based learning or information literacy)? Is it how the library program is organized or scheduled (fixed, flexed or mixed scheduling)? Or, is it how the school librarian works (e.g. providing leadership or collaborating across disciplines and grade levels with other teachers)?

Theoretical Ambiguity

Another unanswered question is how do these many different variables work together to effect student achievement? What is the underlying theory? Many studies assume that a school library that is well-resourced will help students to learn more, and this will result in students’ higher performance on tests. But what types of resources equate to “well-resourced”? Is just having a school librarian enough?

Will a teacher, placed in the position as a school librarian (a common occurrence in many schools) do just as well as a certified school librarian? We know that school librarians deal with reading and that reading has some relationship to learning. But does having twice as many books (or other resources) lead to twice as much learning? Do higher skilled school librarians have more of an effect on student achievement (as in Nationally Board Certified school librarians, and/or school librarians with multiple degrees beyond their certification area)? Do school librarians who remain working in the same school for a longer period of time result in students with higher achievement, and if so, why? Because of this theoretical ambiguity, many school library studies tend to concentrate on inventorying all possible library related resources and activity factors and then try to relate each of those factors to student achievement without any clear theoretical basis for their priority, inclusion or interplay.

The questions listed above, that arise within the context of practice, indicate the current knowledge deficit within the field of school library research and should serve as the basis for future investigations. Haynes suggests that it is important to identify “where the boundary between current knowledge and ignorance lies” (2006, p. 882). Farrugia et al. (2010) suggest the challenge is to develop appropriate research questions to determine “which uncertainties could or should be studied and also rationalizing the need for their investigation” (p. 278). This gap between research and practice has been noted as problematic in other educational disciplines as well, e.g., special education (Cook et al., 2015).

A Field Drowning in Descriptive Data

Second, many of the school library impact studies were exclusively surveys or inventories, and were, therefore, just the assemblage of purely descriptive data of what libraries have or correlational lists of library variables and student achievement. As noted earlier there was a small subset of 27 studies that reflected either a quasi-experimental or an experimental research design. That is, there was no attempt to examine the results of a treatment group relative to a similar control or comparison group in the majority of this research or to try to control for alternative explanations of effects. As noted above, there were only three randomized control trials (RCTs) across all the research studies, and only 24 that used quasi-experimental designs with the use of covariates to account for alternative factors that could influence the outcome. Discounting the three RCT studies using experimental data, of all the 77 other studies with data (of which 44 were in a peer-reviewed publication) reflected the use of observational data. In addition, all of the studies save for a recent study by the authors (Radlick & Stefl-Mabry, 2015) examined a single point in time, rather than using longitudinal data. That means that none of the studies except for one exception, actually examined change in student achievement. All of the studies were a single snapshot at one point in time without controlling for prior achievement. Research designs also reflected a struggle with the size and nature of the samples used in analysis of treatment and control conditions, with many of the studies other than the quasi-experimental design studies, using either small numbers of schools, and/or potentially very unrepresentative samples due to selection bias due to differential responses from subjects. Finally, none of the studies used research designs that addressed the hierarchical, multilevel nature of schools, with classrooms within school which are in turn, within districts and regions.

Unit of Analysis: The Multifaceted Role of a School Librarian

A persistent issue that was almost never discussed was explaining the nature of the school librarian’s work relative to students, classrooms, and schools and how a school librarian’s effect operated across an entire school building. This unit-of-analysis challenge (focusing on school buildings) compounds this problem because the school librarian operates at the school level, so by examining the effect of a single librarian across a single class or a group of classes provides little explanatory power while at the same time requiring a large number of schools to conduct any reasonable analysis. Most of the state-level designs examined the effects of a single school librarian on the school itself, since that is the level at which school librarians typically are allocated and operate. A few studies looked at much smaller groups of students (classroom or grade level) within the school who might be affected by a single school librarian’s interventions. However, designs that examined smaller groups of students within the school building concentrated on very specific programmatic activities or strategies of the school librarian, while building-level designs used a much wider range of school library/librarian resource and activity

variables. The problem of small sample sizes in some studies, providing low statistical power to identify program effects, is a major issue in many of the studies.

General Measurement Challenges

Third, the majority of school library studies suffer from general measurement challenges. Many of the variables, including both school library-related and in a few studies, even student outcomes, reflected solely perceptual measures gathered via questionnaires. Some studies did use formal quantitative survey results of school library resources and activities. As noted above, designs were predominantly non-experimental, so data were at best observational. Even when more formal outcome measures (tests) were used for measuring student achievement, the problem associated with measuring student achievement, including reliability and validity of these various standardized tests, was not discussed. Given a typical school-level analysis, it is critical that assessments are common across large numbers of schools such as in statewide assessments or national standardized assessments in order to assure large enough samples for analysis. With statewide assessments, we have a common proxy measurement of each school's effectiveness that has been generally accepted, albeit with increasing reservations, in the educational community. By using standardized assessments educators are faced with a significant challenge in matching the results of general, once-a-year, high-stakes tests with specific interventions such as school library resources. What is the linkage between a school's curriculum, specific curricular-focused interventions by the school librarian, and a standardized reading assessment? Finally, in some states such as New York, there are high numbers of students opting out of the statewide assessments, and their opting out potentially may reflect disproportional numbers of certain types of students (i.e. high performers or low performers), and thereby make the assessment results less representative of the entire population. In addition, given the unit of analysis problem raised above, we have to address how we are going to represent the assessment results of an entire school building in a single score, especially if the assessment measure is not vertically scaled across grades. Conversely there is the problem of what we use if standardized tests are discounted. How can we argue for a reliable and valid measure of achievement across schools if the measure is some sort of locally developed instrument without any clear psychometric data supporting its reliability or validity?

Problematic Statistical Analysis

Fourth, most of the school library studies suffer from significant statistical analysis problems. In large measure driven by the research designs, most studies used simple descriptive statistics to simply highlight what was taking place in libraries, or used univariate correlational analysis to provide evidence of how the long list of school library variables were related to (positively correlated with) student achievement. Differences in the mean values of variables were generally compared using either multiple t-tests or ANOVA. Only 24 of the total studies used some methodology such as multiple regression to control for covariates that could also explain effects on student achievement. Most of these focused on student economic disadvantage as the primary covariate. As noted above, only the authors' work (Radlick & Stefl-Mabry, 2015) controlled for prior year achievement, along with a number of other variables. The use of hierarchical linear modeling, propensity score matching and other techniques have not been used in the school library research literature, although, as will be discussed below, the authors have modeled student achievement using these techniques. In addition, as Gelman (2016) has written about extensively, too much research suffers from a focus on null hypothesis statistical testing (NHST) that falls into what he calls the garden of forking paths or p-hacking problem—essentially “fishing expeditions” focusing solely on finding statistical significance (p.1). This generally suffers from the problem of interpreting a p value as the size of an effect, rather than focusing on more appropriate effect size measures of program impact within the context of general “noise or uncertainty.” In addition, if we examine many of the school library reports and studies of the effects of school libraries on student achievement, we will see hundreds of statistical tests of statistical significance in each individual study, without any adjustments for multiple comparisons. This fundamental statistical error invalidates many, if not most of the positive results published in the school library literature. This is true because the lack of statistical adjustment for the number of tests increases the significance level required, thereby decreasing the number of significant results that would be found to be statistically significant if adjusted. As Gelman explains, most studies

also suffer from high researcher degrees of freedom, allowing the researcher wide discretion in the selection of different measures and calculations that can easily influence results toward what would be considered a positive outcome by the researcher. By running hundreds of statistical tests until a significant effect is found, and by coupling those tests with a highly selective use of the specific measures that can reflect positively relative to the outcome desired, positive results can be generated, albeit at the expense of a more balanced and realistic view of the effects of school library. Additionally, there were only a few of the regression studies that reported actual effect sizes or variance attributed to school library factors in their models making it difficult to discern the full nature of the analyses, or the specific contribution of the school library variables.

Lack of Replication Studies

Fifth, none of the studies except for the authors' (discussed below) fully replicate studies across multiple outcome measures, multiple populations, or multiple years. Each study is a one-off event. However replication is the essence of good research science, and the commonly cited replication crisis in science reflects the reproducibility problem that exists with so much research, not only in school libraries but across disciplines of education (Makel et al., 2016), psychology (Gilbert, King, Pettigrew, & Wilson, 2016; Solomon, 2016) and even biological sciences (Lewis, Breeze, Charlesworth, Maclaren, & Cooper, 2016; Open Science, 2015; Peng, 2015). However, replication helps to ensure that positive findings from an individual study are reproducible are not merely the result of error, bias, or chance (Cook et al., 2015; Makel et al., 2016).

Hidden Issue of Researcher Bias

Sixth, the issue of research advocacy in support of school librarians has been a hidden issue for decades. It is worth noting, that despite hundreds of studies examining the effect of school libraries on student achievement over the past twenty-five years, not one study has ever challenged or neutralized the notion that school libraries positively impact student achievement. Almost all research related to the effects of school libraries and librarians on student achievement has been conducted by individuals who were intent on justifying claims of school library effects and who publish these results in school library specific publications. In most cases, they were not researchers but school library or library practitioners. In many cases the need to establish an evidentiary basis to support and preserve school librarian funding has been goal of researchers when they initiated the research. As discussed above, there are multiple analytic methodologies, with many inherent weaknesses, that were used to accomplish school library advocacy, however unintended. This is not to say that this research was necessary all bad, but that its often-unarticulated intent was to identify weak results that were positive, and therefore could be used to advocate for school library positions, more staffing and more resources. In fact, the majority of research, as we have shown here and in prior research (Stefl-Mabry, et al, 2016), has been published in journals and other publications associated with the school library field. This means that results that were either not significant, or counter to the predominate positive message of the school library effect on student achievement probably were never published since advocacy publications typically focus exclusively on positive results to support the effective role of the school librarian. This publication bias means that counter instances of researchers reflecting either no impact or negative impact never reach the light of publication. Of course, this leaves practitioners with the misperception that there is strong evidence supporting the role of the school librarian when, in fact, it has been selectively focused on only positive results. "Low-power research designs combined with publication bias favoring positive results together produce a literature with upwardly biased effect sizes" (Open Science Collaboration, 2015, p. 5).

The Difficulty of Conducting Causal Research

Seventh, despite a newfound interest in identifying causal linkages between school libraries and student achievement, as reflected in the Causality: School Libraries and Student Success II (American Association of School Librarians, 2014) initiative by the American Association of School Librarians, proving causal relationships is extremely difficult (American Association of School Librarians, 2014). Over the past decades, there has been extensive debate both within and across disciplines, all struggling to show causality using a wide range of different methodologies. Some researchers argue that observational

data can never show causal relations, and suggest that researchers use experimental data from randomized control trials (RCTs) to assure that both an experimental and a control group are comparable. Others argue that quasi-experimental design studies can demonstrate equivalence between the intervention group and the control. More complex analytic techniques such as structural equation modeling (SEM) provide a way to model and test complex relationships and causal theories going beyond more regression techniques. In addition, quasi-experimental design statistical techniques such as regression discontinuity, propensity score matching, instrumental variables or two staged least squares regression can also provide alternative techniques to demonstrate causal explanations outside of a RCT. Well-designed and well implemented RCTs do allow researchers to better address some of the key challenges such as the hidden or omitted variable problem (which underlies the adage that correlation does not mean causation) which is a serious problem in a non-RCT study. A control group established via random assignment offers a strong counterfactual to show what would take place with the treatment group if there had been no treatment. However, RCTs can often suffer from external validity challenges as noted above due to small sizes required to do practical research. Also, RCTs, without a strong base of theory, do not necessarily address the timing problem (what happens before something and what happens after something, the representative sample size problem (that is assuring a large, representative sample for a study to assure statistical power), the problem of missing data, and finally the problem of measurement error. Pogrow (2017) has detailed the weaknesses inherent in using an RCT-focus such as that by the What Works Clearinghouse. As he illustrates, emphasizing effect size still does not necessarily provide a clear identification of an effective practice. Key to any successful causal or predictive model is how successfully it can be replicated in subsequent populations and datasets with the same results. No single study can provide the certainty that allows us to say that one variable causes another because the possibilities for problems and errors are manifold.

OUR RESEARCH

We embarked on our three-year IMLS research grant (2015-18) with great excitement that we could push the boundaries of school library research on the effects of school librarians on student achievement. Educational researchers are focused on designing studies and measuring changes in students, as well as changes in education practices at the individual student, classroom and school levels. As seasoned educational researchers, we came to this project with our broad experience in research, program assessment and program evaluation. We set out a multi-layered approach, starting with a comprehensive review of all the research literature linking school libraries with student achievement, as well as an analysis and statistical modeling approach using a large-scale, state-level data set reflecting all schools in New York State. We were not prepared for what we found. As discussed in the prior sections of this paper, the state-of-the art relative to school library impact studies leaves much to be desired, with only 24 studies out of nearly 300 actually attempting to address demographic or other factors that would result in non-comparable schools with libraries and those without school libraries. In addition, the body of research suffers, as noted from many significant weaknesses in design and analysis.

Although we would have liked to have undertaken an experimental design with randomization as a part of our IMLS research, we struggled with how to implement a RCT within the context of New York State Schools that would allow us to answer the question of whether school libraries/librarians increase student achievement. The challenges of an RCT design included the question of whether we should focus on a very small number of school libraries and try to examine the results of some library intervention relative to the performance of a group of students in the school, as compared with students in the school who did not get the intervention. To do this we would need agreement from the principal, teachers, librarian, parents and students to randomly select students for the intervention, assuming that the intervention could be implemented within the normal flow of students into and out of the library. Key to this experimentation would be the inequitable distribution of interventions that would be inherent in random assignment. In addition, we would need to obtain achievement and other data for individual students in order to complete the analyses. The number of students and number of schools involved in this kind of design would limit the statistical power that would allow us to discern effects, as well as raise

Results of the first year from the structural equation models, which controlled for many other variables, revealed that school librarians were shown to have a statistically significant impact on student achievement in English Language Arts (both the 2012-2013 ELA Performance Index and the Change in ELA Performance Index from 2011-2012 to 2012-2013). While the school librarian effect on school ELA academic performance was statistically significant, the effect size was relatively small and, not unexpectedly, explained a relatively small part of the variance in total academic achievement (1% of variance explained) as compared with some of the other variables in the models. The path coefficients from the SEM models showed that there were other factors that have a much greater impact on the outcome measures in both ELA and math, such as prior academic performance and poverty, which had the highest effects in the model.

However, despite replicating similar SEM models for years 2013-2014 and 2014-2015 those first results have not been duplicated. In fact, at this point in time we have found little evidence for a school library or school librarian effect. As noted, we have the possibility for multiple replications across grade level groups, subject areas and populations. We would expect that any effect would be consistent, otherwise, there is a high likelihood that a positive result is just a random event. A single positive result does not mean that there is a story there to be told.

Possible Next Steps

We continue our probing both across the large data sets, as well as across the survey data in conjunction with the large data set which can be matched based on school identifiers. Based on our initial research we know that each of the five-school year set of school library data (which is comprised of 45 variables reflecting staffing, operation, print and electronic resources, collaboration and professional development) can account for a maximum of 16% of the variance in student achievement in the building. This means in models with all 45 school library variables, but that have no other covariates, the adjusted R^2 (explained variance) is 16%. But as we know this variance is shared with a number of other things and when those other variables are entered into the model (variables like prior year achievement in that subject area/grade level, demographics, economic disadvantage etc.) the percentage of variance accounted for by school library variables decreases typically to less than 2% and in some cases, it moves to essentially zero. In most cases, this school library contribution is also not statistically significant. We will continue to run different models to attempt to identify significant effects, including exploring subsets of data that reflect only the respondents from the web survey. It is interesting that the handful of the 27 quasi-experimental design studies that did actually report the percentage of variance in their models explained by school librarian/libraries after controlling for admittedly a differing and much smaller set of socioeconomic variables ranged from 21% in a 1994 study (Colorado, 1994), to 2.5% in the Iowa state study (2002) with New Mexico (2002) at 7.9% and Missouri at 11%. One of the more rigorous studies, Wisconsin by Smith (2006) showed a range of 3.2% to 3.4% of the variance explained by school library variables after controlling for other covariates. How do we interpret this? Is the influence of school libraries/librarians lessening over the years? Some evidence for this is reflected in our own results where we have seen a few positive school library effects in the models for the early years (2012-13) (as described above) but not in later years. Are our models allowing us to more appropriately control for variables that influence student achievement? Since we are applying the same models (same variables) in each year, it would not seem to be that this is the case. But we are still looking.

We are also starting to build models with latent variables reflecting key school and library characteristics so that we can use these latent variables in our structural equation models to try to better identify causal effects. The web survey results, while reflecting only a subset of the larger statewide data set, may allow us a representative subset of schools to build models that can help to identify not only resources but actual activities in the library that may be having an effect on student achievement.

Also, because we have been focusing on modeling the mean achievement level using the statistical analysis techniques to date (e.g. multiple regression and maximum likelihood estimators), we may be missing differential effects of school libraries for higher and lower performing schools. Therefore, we will also be using quantile regression models to model the lower quantile results (5th, 10th and 25th percentile) and the higher quantile results (75th, 90th and 95th percentile). The possibility of using a more

longitudinal panel design is also being explored. In short, we are trying to fit the best analytic techniques with the data and questions we have to answer. We would like to be able to identify specific library/librarian effects as a part of our research. But we know how hard this is going to be, given the design and analysis challenges we have experienced already, and outlined here. Another area to explore is the issue of interaction effects between variables. We have had some discussion that this should be pursued further in our research, and it is something we will be looking at in the process of data model considering. Of course, with so many variables, even the number of potential first order interaction effects is huge.

CONCLUSION

As the results of this study indicate, the field of school librarianship research needs to identify effective practices on the basis of multiple, high quality studies that employ experimental research designs in order to demonstrate robust effects on student learning. Research is difficult to conduct in real world educational settings. We suggest that anyone who is engaged in school library research should look long and hard at their measures, the data they collect, their design and their analyses. We recommend alternatives to address weaknesses in research design and analysis and specify a better understanding of the necessary conditions for causal inference in a quantitative study. Such conditions include:

- correlation between the causal agent and the outcome,
- necessity of the cause temporally preceding the outcome
- eliminating possible alternative explanations for any relationship between causes and outcomes (Shadish, Cook, & Campbell, 2002).

The third point is the most challenging, since we can never be certain in our quasi-experimental designs that we are controlling for all the critical alternative variables that could influence the outcome. In addition, the specific questions to be answered by the research need to be posed upfront before the research starts. It is critical that studies should be replicated at least once, and ideally multiple times, to assure the same results before it is accepted for publication. “Replication is the cornerstone of scientific research” and reliable findings from independent investigators is the primary way in which scientific evidence accrues for or against a hypothesis (Peng, 2015, p. 31). This is why multiple replications are critical to confirm causal relationships and prevent us from focusing on random positive effects. We recommend an interdisciplinary approach, with researchers in other disciplines, to encourage knowledge sharing amongst researchers from diverse fields. After all, school librarians collaborate with teachers in educational settings, school library researchers should do the same and engage with non-library researchers to answer questions about how students learn. We suggest working in conjunction with school library practitioners to identify and establish a stronger theoretical base for school library research. We also encourage school library publications to be more critical in their selection of research published and also consider publishing pre-registered studies, along with results, regardless of whether they are positive or not. Only by undertaking a more critical and rigorous approach to school library research will the quality and value of this research be advanced.

REFERENCES

- American Association of School Librarians. (2014). *Causality: School Libraries and Student Success (CLASS). White paper.* Retrieved from <http://libproxy.albany.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED561868&site=ehost-live>
- Anderson, C. A., Leahy, M. J., DelValle, R., Sherman, S., & Tansey, T. N. (2014). Methodological application of multiple case study design using modified consensual qualitative research (CQR) analysis to identify best practices and organizational factors in the public rehabilitation program. *Journal of Vocational Rehabilitation, 41*(2), 87-98. doi:10.3233/JVR-140709
- Barrett, L. (2010). Effective school libraries: Evidence of impact on student achievement. *The School Librarian, 58*(3), 136-139.

- Borkum, E., He, F., & Linden, L., L. (2012). *The effects of school libraries on language skills evidence from a randomized controlled trial in India*. NBER Working Paper Series. Retrieved from <http://www.nber.org/papers/w18183>
- Clyde, L. A. (2004). Research in school librarianship 1991-2000: Australia in an international setting. *The Australian Library Journal*(2), 181-199.
- Cook, B. G., Buysse, V., Klingner, J., Landrum, T. J., McWilliam, R. A., Tankersley, M., & Test, D. W. (2015). CEC's standards for classifying the evidence base of practices in special education. *Remedial and Special Education, 36*(4), 220-234.
- Farrugia, P., Petrisor, B. A., Farrokhyar, F., & Bhandari, M. (2010). Research questions, hypotheses and objectives. *Canadian Journal of Surgery*(4), 278-281.
- Gelman, A. (2016). *Reaction to ASA's statement on p-values*. Retrieved from http://www.stat.columbia.edu/~gelman/research/published/asa_pvalues.pdf
- Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on "Estimating the reproducibility of psychological science". *Science (New York, N.Y.)*, *351*(6277), 1037. doi:10.1126/science.aad7243
- Haynes, B. R. (2006). Forming research questions. *Journal of Clinical Epidemiology, 59*(9), 881-886.
- Hill, C. E. (2012). *Consensual qualitative research: A practical resource for investigating social science phenomena*: Washington, DC : American Psychological Association.
- Hill, C. E., Knox, S., Thompson, B. J., Williams, E. N., Hess, S. A., & Ladany, N. (2005). Consensual qualitative research: An update. *Journal of Counseling Psychology, 52*(2), 196-205.
- Lance, K. C. (1994). The impact of school library media centers on academic achievement. *School Library Media Annual (SLMA), 12*, 188-197.
- Lance, K. C. (2001). Proof of the power: Quality library media programs affect academic achievement. *Multimedia Schools, 8*(4), 14-16, 18, 20.
- Lance, K. C., Rodney, M. J., & Hamilton-Pennell, C. (2003). *How school libraries improve outcomes for children: the New Mexico study*. Santa Fe, N.M.: Hi Willow Research and Publishing for New Mexico State Library.
- Lance, K. C., Rodney, M. J., & Hamilton-Pennell, C. (2005). *Powerful libraries make powerful learners: The Illinois study*. Canton, ILL: School Library Media Association.
- Lance, K. C., Rodney, M. J., & Russell, B. (2007). *How students, teachers & principals benefit from strong school libraries: The Indiana study*. Indianapolis, IN: Association for Indiana Media Educators.
- Lance, K. C., & Schwarz, B. (2012). *How Pennsylvania school libraries pay off: Investments in student achievement and academic standards*. Retrieved from Louisville, Colorado: <http://files.eric.ed.gov/fulltext/ED543418.pdf>
- Lewis, J., Breeze, C. E., Charlesworth, J., Maclaren, O. J., & Cooper, J. (2016). Where next for the reproducibility agenda in computational biology? *BMC Systems Biology, 10*, 1-10.
- Makel, M. C., Plucker, J. A., Freeman, J., Lombardi, A., Simonsen, B., & Coyne, M. (2016). Replication of special education research: Necessary but far too rare. *Remedial and Special Education, 37*(4), 205-212.
- Miles, M. B., & Huberman, M., A. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage Publications.
- Morris, R. J., & Cahill, M. (2017). A Study of how we study: Methodologies of school library research 2007 through July 2015. *School Library Media Research, 20*.
- Open Science, C. (2015). Psychology. estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716. doi:10.1126/science.aac4716
- Open Science Collaboration. (2015). Psychology. Estimating the reproducibility of psychological science. *Science (New York, N.Y.)*, *349*(6251), aac4716. doi:10.1126/science.aac4716
- Peng, R. (2015). The reproducibility crisis in science: A statistical counterattack. *Significance, 12*(3), 30-32.

- Pogrow, S. (2017). The failure of the U.S. education research establishment to identify effective practices: Beware effective practices policies. *Education Policy Analysis Archives*, 25, 1-22.
- Radlick, M., & Stefl-Mabry, J. (2015). *Finally – Convincing evidence for the impact of school librarians?* Paper presented at the American Educational Research Association, Chicago, IL, April 16-20.
- Scholastic Research & Results. (2008). *School libraries work!* Retrieved from <http://www.p12.nysed.gov/technology/library/libnatandinter.html>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Smith, E. G. (2006). *Student Learning Through Wisconsin School Library Media Centers: Case Study Report*. Retrieved from Madison, Wisconsin: <http://dpi.wi.gov/sites/default/files/imce/imt/pdf/finalcasestudy.pdf>
- Solomon, B. G. (2016). The implications of 'estimating the reproducibility of psychological science' for school psychology research. *Science*, 9.
- Stefl-Mabry, J., Radlick, M., Armbruster, D., & Keller, Y. (2016). *Breaking down Information silos: Sharing decades of school library research with educational researchers*. Paper presented at the 2016 American Educational Research Association Annual Meeting, Washington, D.C. , April 8-12.
- Subramaniam, M. (2015). New territory for school library research *Knowledge Quest*, 43(3), 16-19.
- What Works Clearinghouse. (n.d.). *What Works Clearinghouse™ procedures and standards handbook version 3.0* (pp. 91). Washington, DC: Institute of Education Sciences (IES).