# The Southeastern Librarian

Volume 69 | Issue 1 Article 3

Spring 5-7-2021

# Citation Analysis: Availability of Student Research Poster **Citations**

Sharon Holderman Tennessee Tech University, sholderman@tntech.edu

Stuart Gaetjens Tennessee Tech University, sgaetjens@tntech.edu

Follow this and additional works at: https://digitalcommons.kennesaw.edu/seln



Part of the Library and Information Science Commons

# **Recommended Citation**

Holderman, Sharon and Gaetjens, Stuart (2021) "Citation Analysis: Availability of Student Research Poster Citations," The Southeastern Librarian: Vol. 69: Iss. 1, Article 3.

Available at: https://digitalcommons.kennesaw.edu/seln/vol69/iss1/3

This Article is brought to you for free and open access by DigitalCommons@Kennesaw State University. It has been accepted for inclusion in The Southeastern Librarian by an authorized editor of DigitalCommons@Kennesaw State University. For more information, please contact digitalcommons@kennesaw.edu.





# Citation Analysis: Availability of Student Research Poster Citations

Sharon Holderman<sup>a</sup> (sholderman@tntech.edu) and Stuart Gaetjens<sup>a</sup> (sgaetjens@tntech.edu) <sup>a</sup>Tennessee Tech University, Cookeville, Tennessee, USA

#### ABSTRACT

This citation analysis examines the availability of 1313 citations from student research posters across multiple disciplines at a mid-sized university. The study focuses on whether the cited materials are available via the library or freely online. Results indicate only one-fifth of citations were from library proprietary sources while two-thirds were freely available online. Resource age and discipline also influence availability. This study informs collection development decisions and instruction services and helps define the library's role of providing information and facilitating information access.

#### **KEYWORDS**

citation analysis, collection development, student research

## **INTRODUCTION**

Students' academic needs are a moving target, which makes it challenging to develop a library collection for them. The academic demands of the institution drive the students' need for library resources. Students receive assignments from faculty requiring resources with different restrictions like topic, quantity, resource type, and age. These assignment factors make it difficult to anticipate students' library collection needs accurately.

One of the most influential factors in collection development is the wealth of information freely available online. While some of this free information is not suitable for student use, much of it is appropriate; there are open-source journals rigorous enough for research papers as well as full-text, peer-reviewed articles posted by authors on sites like ResearchGate, a professional network to share and discover research. Even sites like SciHub, which offer access to content while violating copyright, do offer students a way to obtain scholarly resource material for their academic work. Ideally, libraries also offer a useful, relevant collection for students to use. One way to help create a useful collection is to look at what resources students have used in the past. Circulation and electronic resource usage statistics help, but there is still more to learn. Citation analyses offer insights into which resources students are using, which can then be cross-referenced with library holdings to help evaluate the collection.

Flat or declining academic library budgets are also a major factor in collection development. Across the nation, state funding for public colleges and universities declined by \$9 million between 2008 and 2017, and an ACRL survey found 19% of libraries had decreased funding while 60% had flat budgets between 2010 and 2015 (Bosch et al., 2018). Even flat budgets are problematic when confronted with an average 6% increase in subscription prices every year (Bosch et al., 2018). These conditions create pressure for libraries to spend wisely while trying to satisfy patrons' information needs.

# LITERATURE REVIEW

Citation analyses study what resources students are using and typically include a breakdown by resource type (Sherriff, 2010). These analyses are most helpful for collection development interpretations when they also search for those citations in the library. There are a number of citation analyses, outlined in Table 1, that offer insight into resource type and library availability. University of Notre Dame libraries conducted one of the largest analyses by reviewing over 27,000 unique citations in all doctoral dissertations submitted between 2005 and 2007; 55% of citations were journals and 37% were books (Kayongo & Helm, 2012). The library paid for either the print or electronic version of 67% of the citations, which means those doctoral students could use the library to access their materials about two-

thirds of the time (Kayongo & Helm, 2012). These results were somewhat similar to the study at University of Texas at Austin, which focused on engineering and educational psychology dissertations between 1997 and 2002 (Fuchs et al., 2006). Of their 3,120 citations, 28% were books, 51% were journals, and overall the library owned 78% of all the citations (Fuchs et al., 2006).

Southern Connecticut State University studied history theses written between 1998 and 2008 and found the resource type ratios to be reversed, and 59% of the 3,498 citations were books while only 24% were journals, which indicates the most popular type of resource can vary based on the research discipline (Sherriff, 2010). One study of graduate students' citations across disciplines found a more traditional breakdown of resource type but much higher library ownership rates. A University of Georgia study reviewed 3,363 theses and dissertation citations across multiple disciplines from 1991 and 2001, which revealed the library owned 85% of cited books, and 92% of cited journals (Smith, 2003). This high rate of library ownership suggests a strong connection between their graduate students' research and use of the library collections.

This high rate of ownership was also evident in the University of South Alabama's study of 2,301 citations from undergraduate honors theses (Wilson, 2012). That library owned 60% of the book citations and 91% of the journal citations (Wilson, 2012). These ownership numbers are similar to Eastern Illinois University's study of undergraduate writing portfolios where their library owned 55% of the cited books and 80% of the cited journals (Knight-Davis & Sung, 2008). These high rates of ownership do not necessarily mean collections truly satisfy students' needs because it is possible students are using the collection as their main resource instead of looking elsewhere or using interlibrary loan.

These studies, like most citation analyses, focus on library availability related to ownership by paid access. Typically, they do not look at whether materials are freely available online. Only Arcadia University used that parameter in their study of 1,347 citations from 2012-2013 senior theses in the social sciences (Kohn & Gordon, 2014). They found 34% of cited books were available via the library, and 10% were available freely online, which means 56% of the books were obtained elsewhere (Kohn & Gordon, 2014). Cited articles were available via the library 78% of the time and available freely online 6% of the time, which means students obtained 16% of the articles elsewhere (Kohn & Gordon, 2014). Items were only searched for free online access if they were not available via the library, so there was no overlap analysis (K. Kohn, personal communication, March 21, 2019).

Studying citations for free online access acknowledges students may be obtaining sources from the internet, which is highly likely. Students are driven by a need to find sources based on their content, and they will access material based on their preferences for searching and ease of access. How they fill their need for information is most likely driven by faculty guidelines related to resource type, reliability, and location. Because libraries do not always know what these faculty guidelines are, comparing library availability with free online access offers a more diverse perspective on resource accessibility.

Table 1: Literature Review Overview							
Authors	University	Citations	% Book	% of Books	% Journal	% of Journals	% of Citations
			Citations	Owned	Citations	Owned	Owned
Fuchs, Thomsen, Bias, Davis	University of Texas at Austin	3120	28%		51%		78%
Kayongo, Helm	University of Notre Dame	39,106	37%		55%		67%
Knight_	Factorn						

Table 1. Literature Review Overview

#### **BACKGROUND**

Tennessee Tech University is a state university with approximately 10,000 students located in Cookeville, TN. The library has a collections budget of approximately \$735,000 for both physical and electronic resources in the 2018-19 fiscal year, which is nearly the same dollar amount allocated in 1988-1989 (Bates, 2018). The collections budget generally increased from 1988 until it hit its peak in 2009-2010 at \$1.34 million, which means the library has experienced approximately \$600,000 in budget reductions over the last 10 years (Bates, 2018). Due to inflation and enrollment increases, the library cannot serve students with materials the same way as 1988 even though the budgets are the same; in 1988-1989, the library collection budget allowed for spending \$189 per student, but in 2018-2019 the budget allowed for \$72 per student (Bates, 2018). This ranks Tennessee Tech's library as 111 out of 115 schools in collection spending per student compared to other academic institutions with the same Carnegie Classification of Doctoral University: High Research Activity (Bates, 2019). This classification is also known as R2 and means Tennessee Tech awarded at least 20 research/scholarship doctoral degrees and had over \$5 million in research expenditures in the update year (The Carnegie Classification of Institutions of Higher Education, n.d.). These budget struggles reinforce how important it is for the library to be responsible when spending collection money.

The goal of the collection is to help the campus community conduct research and complete course assignments. The library reviews circulation statistics for physical items, but this does not include print journals because they do not circulate. The library also reviews usage statistics for online subscriptions prior to renewal to ensure the resource is being used enough to warrant spending the money. Budget cuts over the last 10 years have required repeated subscription cancellations, which are decided based on price, usage statistics, and relationships to academic programs. The problem with relying heavily on usage statistics is although they prove a title or database is being used, it does not indicate whether that database is the perfect solution to students' research needs. It is possible students use a resource because it is there and available but would actually prefer something else if the library offered it or if they had enough time to request it. This type of information substitution would have little impact in introductory class assignments requiring five sources on a broad topic, but it is more relevant for advanced student research with more specific information needs.

Tennessee Tech students gather research from a variety of sources while participating in the university's annual Research and Creative Inquiry Day. This event began in 2007 with 67 posters and grew to over 200 posters by 2018 (Office of Research, 2019). The event showcases research by undergraduate and graduate students from all departments and includes awards for research posters. Beginning in 2017, students participating in Research Day could choose to publish their abstracts and posters in an online journal (Proceedings of student research and creative inquiry day, 2018). There are no reference guidelines for the posters; students are not required to include references on their poster. However, most do include references, and the majority are scholarly in nature. Students work with faculty on their research posters, so it is possible faculty are offering guidance on resource use, but there are no official requirements regarding resources as part of Research Day. The lack of official resource guidelines and multi-disciplinary participation makes Research Day posters a good venue for determining which sources student researchers choose to use.

#### **METHODS**

This study began with an interest in students' preferred research sources. Citations were collected from Research Day posters in 2016, 2017, and 2018. Citations from 2016 were obtained from photos taken at the event, which resulted in using 51 of 224 posters with 230 citations. In 2017, there were 276 citations from 54 of 189 posters, which were obtained from photographs and the publication of the proceedings. The photos from 2016 and 2017 were taken as part of promotional efforts, which is why not all posters were included in those years; the photos were taken at random. In 2018, all 218 posters were reviewed because photos were taken of every poster for this analysis, yielding a total of 807 citations. Across all years, the only posters intentionally eliminated from the analysis were those not containing citations.

Each citation was entered into an Excel spreadsheet and was coded for additional data including poster title, poster year, poster discipline, citation year, and whether the resource type was a book, article, website, or other. Articles were identified as having an author and written in an article format regardless of whether it came from a journal, magazine, newspaper, government website, etc. Since not all students used an approved citation method, some citations were incomplete and were not used in the analysis. Duplicate citations were kept if they were used for different posters because they were legitimately used twice as a source.

Citations were also coded for availability in four categories: library print, library electronic, freely available online, or undetermined. Undetermined meant the material could not be found in the other three categories, so it was unclear how the student obtained the resource. Items with overlapping availability were marked in all applicable categories. Library availability for books was determined by searching the online catalog, which found both print and eBooks. The library's discovery tool and eJournal search were used to determine the library availability of articles. Articles were marked with library availability if they were in proprietary databases and journals; free databases like PubMed Central, although searchable in the library's discovery tool, were coded as freely available online and not library available. The balance of this paper will focus on the unique proprietary items rather than including the overlapping citations also freely available on the internet. Drawing this line was challenging because Tennessee Tech has access to some materials purchased by other state entities; databases purchased for online degree programs in the state were included as proprietary, but databases free to all residents in the state were considered freely available online. Whether citations were freely available online was determined by searching for items in Google Scholar and Google and clicking through to the resource to verify the full text was actually available. These two search engines were chosen because of their popularity; students use them frequently based on anecdotal evidence from the reference and instruction librarians. Searching for items freely available online was performed off-campus to ensure free access was not confused with vendor-established access through campus IP addresses.

There were two areas of access originally considered for investigation but ultimately eliminated from this analysis. Many proprietary and undetermined articles were freely available through SciHub, but it was unclear whether students would use this, especially since it requires specific searching at that site and not through Google. An informal, email survey of the 2018 poster faculty advisors indicated they do not use SciHub nor do they recommend students use it. Additionally, many of the responses indicated they had never heard of SciHub. Therefore, resources coded as freely available online were found through Google Scholar or Google only.

The second potential consideration was looking at InterLibrary Loan (ILL) data to see if the undetermined resources were acquired there. Data was pulled from ILLiad, Tennessee Tech's ILL vendor, but matching the citations to this data proved too difficult. The ILL data did not include patron names, which made it impossible to know if the ILL request was made by a poster author. There was also often a time delay of one to two years between the ILL request and the poster date, which made it too difficult to verify the ILL request was directly related to the poster. Therefore, ILL usage was not included in the scope of this analysis.

## **RESULTS**

The goal of this analysis was to study the availability of resources in poster citations. As detailed in Table 2, there was overlap between library availability and items freely available online. However, this analysis focused more on items only available through the library versus items freely available online. Results indicated 18% of the resources used were library proprietary, which was approximately the same percentage as the undetermined resources and the resources available both through the library and freely online. Since approximately 18% of the citations were not available online or through the library, students may have obtained materials from ILL. Some undetermined article citations listed authors recognized as Tennessee Tech faculty members, so it is possible students obtained those resources directly from their faculty poster advisors. Students had free, online access to 63% of all citations.

**Table 2:** Availability of Citations

Library proprietary	Both library proprietary and free online	Free online only	Undetermined	Total
242 (18.4%)	234 (17.8%)	598 (45.5%)	239 (18.2%)	1313 (100%)

Students used articles far more often than other resource types as demonstrated in Table 3. Approximately one-third of the book citations were coded as uniquely available through the library, another third were freely available online, and the last third had an undetermined availability. Students used more websites than books for their poster citations. When it came to articles, 21% of the citations were coded as available solely as library proprietary sources while 62% of them were freely available online.

**Table 3:** Resource Type and Availability

	Library proprietary	Both library proprietary and free online	Free online only	Undetermined	Total
Book	50 (32.5%)	9 (5.8%)	39 (25.3%)	56 (36.4%)	154 (100%)
Article	192 (20.7%)	225 (24.3%)	351 (37.9%)	158 (17.1%)	926 (100%)
Website	n/a	n/a	208 (100%)	n/a	208 (100%)
Other	0	0	0	25	25 (100%)

Resource age appears related to library availability as shown in Table 4. The most recent resources had the highest percentage of free online access while the oldest resources were least available freely online. The opposite was true for library availability; the most recent resources had the lowest percentage of library availability, and the oldest resources had the highest. Generally, library availability increased as resource age increased, and free online access decreased as the resource age increased.

**Table 4:** Resource Age and Availability

	Library proprietary	Both library proprietary and free online	Free online only	Undetermined	Total
o-3 years old	47 (10.9%)	55 (12.8%)	252 (58.6%)	76 (17.7%)	430 (100%)
4-6 years old	47 (19.3%)	55 (22.5%)	103 (42.2%)	39 (16.0%)	244 (100%)
7-10 years old	51 (23.9%)	64 (30.0%)	78 (36.6%)	23 (10.8%)	213 (100%)
11-15 years old	26 (19.5%)	25 (18.8%)	60 (45.1%)	22 (16.5%)	133 (100%)
16+ years old	71 (27.8%)	38 (14.9%)	72 (28.2%)	74 (29.0%)	255 (100%)

Lastly, online availability, regardless of library availability, was evaluated by discipline, shown in Table 5. Computer science had the highest percentage of citations freely available online at approximately 93%, and human ecology was the next highest discipline with approximately 73%. Humanities had the lowest at approximately 31%. Unique library availability was highest for humanities at almost 38% followed by education at almost 30%.

**Table 5:** Discipline and Availability

	Library proprietary	Both library proprietary and free online	Free online only	Undetermined	Total
Computer Science	3 (3.3%)	18 (20.0%)	66 (73.3%)	3 (3.3%)	90 (100%)
Education	17 (28.8%)	13 (22.0%)	24 (40.7%)	5 (8.5%)	59 (100%)
Engineering	64 (13.6%)	80 (17.0%)	215 (45.6%)	112 (23.8%)	471 (100%)
Human Ecology	47 (21.2%)	59 (26.6%)	104 (46.8%)	12 (5.4%)	222 (100%)
Humanities	28 (37.8%)	6 (8.1%)	17 (23.0%)	23 (31.1%)	74 (100%)
Sciences	76 (22.2%)	48 (14.0%)	141 (41.2%)	77 (22.5%)	342 (100%)

Note. Only disciplines with 50 or more citations were included in the discipline evaluation.

## **IMPLICATIONS**

The authors were surprised at how many citations were freely available online, which was approximately two-thirds, especially considering that the citations came from Research Day posters and not introductory freshman-level course assignments. Although it is impossible to determine where students obtained the materials that are both library proprietary and freely available online, the reality is that they would not need to use the library to obtain them. It is also unclear whether students were seeking a specific article and happened to find it freely available online, or if they searched online for freely available material and used what they found. Either way, the library collection appeared to be used minimally for Research Day posters, which suggests the collection was not heavily needed since students completed their research and posters successfully.

Articles were cited much more frequently than other resource types, but article availability coded as uniquely library occurred approximately 21% of the time. This low percentage for articles is disappointing, especially considering approximately 80% of Tennessee Tech's library collections budget is spent on access to articles (Gaetjens, 2019). These statistics invite the following questions: can the money paid for articles through database and journal subscriptions be better allocated? Or does Tennessee Tech need to invest more in articles? Would purchasing individual articles on demand be a better way to meet students' needs? Ideally, librarians would review subscription purchases and renewals for overlap with material freely available online. This would be easier with individual journal subscriptions, but even then the overlap would not necessarily justify canceling a subscription.

The number of total website citations exceeded the number of book citations. Further investigation may reveal why websites are a more popular resource type than books. Possible reasons include information timeliness, easier access, and keyword searching. Libraries have often stressed the need for peer-reviewed and scholarly resources for research, but the high use of websites suggests that may not reflect the reality of students' needs. Students have faculty advisors for their Research Day posters, which means that faculty are accepting this frequent website usage if not encouraging it. If valuing and using websites as sources more than books is part of a larger trend for university students, this could suggest less money should be spent on books. Perhaps eBooks are a better option, but those would be library-purchased materials and require library authentication and access. If students are searching for information freely available online, they may not find library-purchased eBooks because of the proprietary access.

This study also suggests the library collection may be out of date, particularly for certain disciplines, because library availability increased as the resource age increased. This observation may not be as problematic as it first appears if libraries see potential in providing older materials in certain disciplines where it might be more useful to students. Then students can access the library's older materials as well as online materials freely available online to have more resources available to them for their research. Books will naturally become older as time passes, but perhaps subscription money could be focused more on backfiles to help provide access to older materials in relevant disciplines.

If libraries consider freely available materials a viable source of information for students, and especially if they rely upon those free materials to complement their collection, they should also help students find online information effectively. Instruction sessions can demonstrate how to best search Google Scholar, ResearchGate, and other relevant sites. Teaching students how to connect Google Scholar to their library's full text can help meet their information needs while minimizing the number of places they need to search.

Scrutinizing the citations by discipline may be valuable input for collection development decisions. For example, over 93% of computer science citations were freely available online. In the interest of spending collection money responsibly, logic suggests that less money should be spent on resources in computer science because so many are easily accessible and free. The only discipline where the library proprietary citations outnumbered those freely available online was in the humanities. This could suggest humanities students purposely seek library materials more, or the humanities materials in the library are more useful, or a combination of both. Another discipline-specific statistic worth considering in collection development is that the engineering, humanities, and science disciplines used citations, which were not available freely online or in the library's collection, approximately 22% to 31% of the time. Although ILL is available to obtain these materials, perhaps collection dollars should be focused in these areas so students have more immediate access to the resources they need.

#### **CONCLUSION**

A citation analysis cannot indicate exactly where students accessed their resources, but it can start valuable conversations related to whether an academic library holds information that is being used by their students. In general, analyses also do not indicate why students use specific resources, so it is unclear whether they are more apt to use what is readily accessible or whether they are intentionally seeking a specific resource. Students' resource choices can be influenced by their search venue, research skills, and availability (Fuchs et al., 2006), which is why this study examined where those resources are available and not whether those resources were the best choice.

One of the strengths of this study is its focus on the authors' own university, which makes the results and implications very relevant to a specific library. The data itself can show trends for other universities, and the study helps demonstrate why it is important to consider how to increase the collection's impact when making budget decisions. This study can also be used as a blueprint for citation analyses for other libraries. The study's focus on Research Day targets students voluntarily participating in research, which is usually more scholarly than freshman introductory papers. Research Day is also interdisciplinary to help address multiple areas of the collection instead of studying papers in one class or major.

This study does have some limitations. First, some disciplines had very few citations over the three-year period, and that lack of data prevents meaningful insight into those areas. Next, the citations from 2016 and 2017 are not all-inclusive like 2018, so not all poster years were represented in the same way. Finally, poster research is not as common as students doing research for class assignments, although some of the posters are tied to class assignments. Therefore, conclusions drawn from this analysis may not apply to the larger student population.

Since citation analyses do not comprehensively address student usage and resource availability, additional studies help offer a more comprehensive picture. Similar studies across various classes and departments would give a more complete picture of various types of research beyond Research Day posters. Conducting an additional analysis on disciplines underrepresented at Research Day would be beneficial, especially where purchases are high-cost or where usage data does not match anecdotal data. Further study on the undetermined access points would be beneficial, which could be done in part by studying ILL data. Adding a student survey element could indicate where students are actually obtaining their resources and not just where the resources are available. Lastly, it is difficult to determine if the high percentage of citations freely available online is due to the low level of library material spending at Tennessee Tech or success in the open access movement. Further research at research universities which

spend more on library materials would help put this observation into a broader context.

In the past, libraries were considered the information gatekeepers for their academic communities. With the increase of high-quality free information on the internet, this role has become untenable. When comparing the results of this study to previous studies, one sees that the portion of useful information materials purchased by the library in this study is lower than in any previous study. That could be because the quantity of high-quality free information from websites, institutional repositories, and from publishers has expanded a great deal over the past 20 years. As libraries transition from gatekeepers to a new role more as interested bystanders, difficult conversations arise about collection development in libraries.

## **REFERENCES**

- Bates, D. (2018). Budget-materials [Unpublished raw data].
- Bates, D. (2019). Data Carnegie [Unpublished raw data].
- Bosch, S., Albee, B., & Henderson, K. (2018). Death by 1,000 cuts: Periodical price survey 2018. *Library Journal*, 143(7), 28-33. https://www.libraryjournal.com/?detailStory=death-1000-cuts-periodicals-price-survey-2018
- The Carnegie Classification of Institutions of Higher Education. (n.d.). *Basic classification*. https://carnegieclassifications.iu.edu/classification\_descriptions/basic.php
- Fuchs, B. E., Thomsen, C. M., Bias, R. G., & Davis, D. G. (2006). Behavioral citation analysis: Toward collection enhancement for users. *College & Research Libraries*, *67*(4), 304-324. https://doi.org/10.5860/crl.67.4.304
- Gaetjens, S. (2019). Monthly materials expenditures [Unpublished raw data].
- Kayongo, J., & Helm, C. (2012). Relevance of library collections for graduate student research: A citation analysis study of doctoral dissertations at Notre Dame. *College & Research Libraries*, 73(1), 47-67. https://doi.org/10.5860/crl-211
- Knight-Davis, S., & Sung, J. S. (2008). Analysis of citations in undergraduate papers. *College & Research Libraries*, 69(5), 447-458. https://doi.org/10.5860/crl.69.5.447
- Kohn, K. C., & Gordon, L. (2014). Citation analysis as a tool for collection development and instruction. *Collection Management*, *39*(4), 275-296. https://doi.org/10.1080/01462679.2014.935904
- Office of Research. (2019). TTU research and creative inquiry day. https://www.tntech.edu/research/research/av/
- Proceedings of student research and creative inquiry day. (2018). https://publish.tntech.edu/index.php/ PSRCI/about
- Sherriff, G. (2010). Information use in history research: A citation analysis of master's level theses. *portal: Libraries and the Academy*, 10(2), 165-183. https://doi.org/10.1353/pla.0.0092
- Smith, E. T. (2003). Assessing collection usefulness: An investigation of library ownership of the resources graduate students use. *College & Research Libraries*, *64*(5), 344-355. https://doi.org/10.5860/crl.64.5.344
- Wilson, E. K. (2012). Citation analysis of undergraduate honors theses. *The Southeastern Librarian, 60* (1), 39-50. https://digitalcommons.kennesaw.edu/seln/vol6o/iss1/7/