

11-17-2008

Dissertation Citations in Organismal Biology at Southern Illinois University at Carbondale: Implications for Collection Development

Jonathan Nabe

Southern Illinois University Carbondale, jnabe@lib.siu.edu

Andrea Imre

Southern Illinois University Carbondale, aimre@lib.siu.edu

Follow this and additional works at: http://opensiuc.lib.siu.edu/morris_articles

 Part of the [Library and Information Science Commons](#)

Published in *Issues in Science and Technology Librarianship*, No. 55, Fall 2008.

Recommended Citation

Nabe, Jonathan and Imre, Andrea. "Dissertation Citations in Organismal Biology at Southern Illinois University at Carbondale: Implications for Collection Development." (Nov 2008).

This Article is brought to you for free and open access by the Morris Library at OpenSIUC. It has been accepted for inclusion in Articles by an authorized administrator of OpenSIUC. For more information, please contact opensiuc@lib.siu.edu.

Dissertation Citations in Organismal Biology at Southern Illinois University at Carbondale: Implications for Collection Development

Abstract

We report on a citation analysis of PhD dissertations in Plant Biology and Zoology at Southern Illinois University Carbondale, undertaken to test the common assumption that scientists favor current research to such an extent that journal backfiles can be de-emphasized in academic library collections. Results demonstrate otherwise. The study is reproducible for any institution, and can help to evaluate 1) the value of electronic journal backfiles and 2) the need to maintain print backfiles.

Introduction

Conventional wisdom says that only the most current materials are useful and in high demand for the sciences. If so, the implications for strained budgets are that journal backfiles for the sciences are a poor investment, and that removing older bound journal volumes to create space and/or reduce shelving expense is a sound collection development decision. That assumption is tested here, using dissertation citations from two science disciplines.

There are many reasons for conducting a local citation study. Two pertinent reasons related to collection development are consideration of the removal of bound print volumes of journals, and consideration of electronic journal backfiles for purchase or lease. This study applies a citation study towards the evaluation of these issues, and demonstrates how similar studies can be applied elsewhere.

Shelving space is expensive, and space itself is often in high demand. Many libraries are under pressure to make difficult decisions regarding older materials sitting on shelves, including print journals. On another front, electronic journal backfile offerings from publishers are commonplace now, with all of the major publishers having some archival package available. In order to assess the value of these backfiles, and/or the need to maintain bound print volumes, some measure of the potential use is essential.

Past and current use of older volumes of journals can be used as a proxy for such an assessment. Two methods are available: shelving studies and citation studies. Shelving studies are useful, in the same way that download reports from journal publisher web sites are, but suffer from some of the same limitations. There is no way to know how the journal was used, nor who used it. In addition, they are painstaking, slow, and highly susceptible to human error and inconsistencies. While useful, a more accurate measurement is desirable.

Citation studies offer reliability, relevance, and reasonable speed. Citation studies of dissertations provide valuable information for the most important, higher profile programs of an institution. Since these programs are likely to be major selling points for the university, as well as primary earners of grants, providing the necessary resources to support them is a high priority for any academic library. Dissertations clearly indicate the needs of graduate students, and also indicate the research specialties of the faculty and departments as a whole. With exceptions for new (or defunct) programs, current and historic data is readily available in an institution's dissertations, though significant processing is required.

One area in which citation studies have not been conducted much is organismal biology, composed of the fields of botany (here called plant biology) and zoology. This study examines

these two disciplines both to fill this gap and because both of these departments at Southern Illinois University Carbondale (SIU-C) offer PhDs. Because of the similarity in research methods and in order to provide a large enough data sample, dissertations from the two departments are examined together, with any disparities between the two described.

Background

SIU-Carbondale is an ARL, Carnegie RU/H university, and has both a Law School and a Medical School. Enrollment is 21,000, including 4,790 graduate and professional students in 63 masters and 29 doctoral programs. Situated on the edge of the Shawnee National Forest, encompassing over 280,000 acres, SIUC is within close distance of numerous unique wilderness areas, including LaRue Pine Hills, which contains over 1,200 species of vascular plants within its 4.5 square miles.

The Department of Plant Biology has sixteen faculty members, and an average enrollment of about twenty PhD students, and twenty Master's students. The Department arranges its programs around three nodes: Ecology, Molecular and Biochemical Physiology, and Systematics and Biodiversity. Its strengths are in forest, grassland, and wetland ecology, plant systematics and evolution, mycology, bryology and lichenology, pollination ecology, plant molecular biology, molecular biology of mineral nutrition and micronutrient uptake, developmental morphology and anatomy, plant stress physiology, and electron microscopy.

The Department of Zoology has twenty-seven faculty, averages thirty-five PhD students and 75 Master's students. The Department provides a broad range of expertise in almost all areas, and is particularly strong in wildlife ecology and management, and fisheries.

Literature Review

Citation studies of dissertations in the sciences are infrequent. What studies have been done usually provide little or no data or analysis on the *age* of citations, focusing instead on material format and specific journal titles. Vallmitjana and Sabaté (2008) noted a median age of nine years for a sample of chemistry dissertations from the Institut Quimic de Sarrià in Spain. Walcott (1994) reported on dissertations from all of the Biological Sciences at SUNY Stony Brook, noting that 50% of the citations she studied were no older than five years, and 80% were no older than ten years. She did, however, note a strong difference between Ecology and Evolution dissertations compared to Molecular Biology, Genetics, and Neurobiology. Brazzeal and Fowler (2005) reported an average citation age of 10.6 years for Master's theses in forestry from Mississippi State University. Kushkowski et al. (2003) studied theses and dissertations from Iowa State and found a mean citation age slightly less than 15 years, but they did not separate theses from dissertations, and used the broad category of "Biological Sciences," which included everything from agronomy to zoology, including bacteriology, genetics, microbiology, etc. Edwards (1999) studied polymer theses and dissertations but did not report on the age of citations; the same is true for Gooden (2001) who studied chemistry dissertations at Ohio State. Williams and Fletcher (2006) reported on engineering masters' theses at Mississippi State, and using the same measurements as Walcott, noted that across all engineering fields, the median age was 7.5 years, and 80% were eighteen years old or less.

Some studies of publications by faculty or by specific journals have looked at citation age more systematically. Salisbury et al. (2007) reported on University of Arkansas Food Science faculty publications. Although not reporting median age, she notes that 44% of cited journal articles

were thirteen years old or less. Ackerson (2001) found that in the top journals in physical chemistry, the median age was between five and seven years, depending on the type of journal (original research vs. review). Musser and Conkling (1996) found the 50% rate of “major scholarly journal” citations in engineering to be eight years old, and 75% were within sixteen years, and concluding that ten to fifteen years of backfiles would be sufficient for most users. Delendick (1990) found that a minimum of 67% and a maximum of 87% of citations to three plant systematics journals were at least twelve years old.

Methods

All dissertations from the Departments of Plant Biology and Zoology from 2003 – 2007 were analyzed. By restricting the study to dissertations, and excluding master’s level theses, resource demand for a higher level of research is uncovered. In addition, restricting studies to dissertations will allow for a smaller pool of documents, making a more manageable project, without being subject to the inaccuracies that sampling may bring. In this case, thirty dissertations were analyzed, seventeen from Zoology and thirteen from Plant Biology. Five years provides a viable window of current research needs, though going back further would be necessary to analyze longer term trends. Dissertations were identified via ProQuest’s Dissertations & Theses and verified against departmental information.

All citations from each dissertation were included for analysis, thus avoiding any sampling error. Again, by limiting the study to dissertations, a more accurate picture is attained. A Perl script was used to separate citation components for automatic importing into an Excel spreadsheet, but this was only possible for full-text searchable pdfs whose citations had been copied into a .txt file, and spaces manually entered between the citations. For pdfs that were not full-text searchable, data for analysis was manually entered into the Excel spreadsheet, including dissertation author, citation date, and a code to indicate the format of the citation: journal, book, proceeding, series, report, map, newspaper, web site, software, or data set.

Results

There were a total of 4,563 citations for these five years of dissertations in Zoology and Plant Biology. 1,450 were from nine 2007 dissertations; 680 from five in 2006; 798 from five in 2005; 1,125 from seven in 2004; and 510 from four in 2003 (See Table 1).

Dissertation Year	Number of Dissertations	Number of Citations	Average Number of Citations	Number of Journal Citations	Average Number of Journal Citations	% of Citations to Journals	% of Citations to Journals Pre-1996
2007	9	1,450	161	1,117	124	77%	47%
2006	5	680	136	496	99	73%	41%
2005	5	798	160	542	108	68%	62%
2004	7	1,125	161	738	105	66%	64%
2003	4	510	128	374	94	73%	77%
Totals	30	4563	152	3267	109	72%	56%

Table 1. Citation Breakdown by Year

To assess the value of electronic backfiles, the percentage of citations to journal articles prior to 1996 was calculated. 1996 was chosen as a conservative date to account for the earliest beginning of online access included with standard online subscriptions. This cutoff date applies to most titles published by commercial publishers including Wiley, Elsevier, Springer as well as such university presses as Oxford. Most publishers in the sciences will require an archival purchase or lease to have access to content before that year. The dissertations from all years show a significant percentage of citations to resources dating before 1996. For citations to journal articles only, 56% (1,821 of 3,267) fall into that category. Of 4,563 total citations, 2,597, or 57%, are pre-1996. Also of note, of citations to monographs, which include books, proceedings, reports, and theses/dissertations, 759 out of 1,219 are pre-1996 (62%).

Results are not skewed by outliers (Figure 1). Fifteen authors' percentage of pre-1996 citations to journal articles fall below the average of 56%, fifteen above. Eight of the thirty authors have above 70% pre-1996 citations. Only one author had below 20% pre-1996 citations percentage, and only six were below 40%.

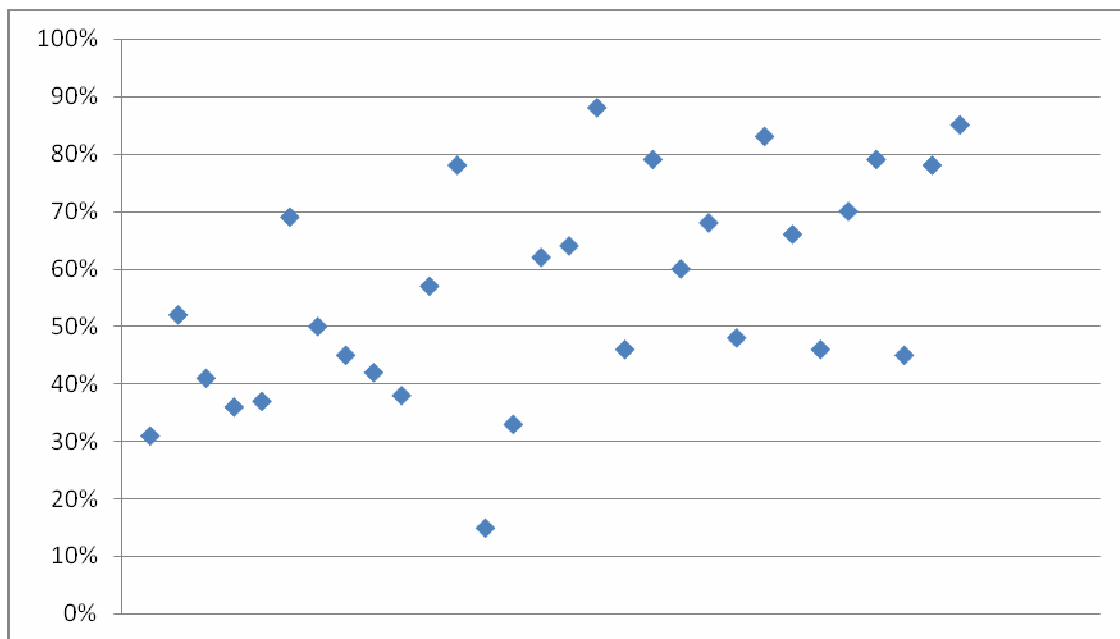


Figure 1. Individual Authors' Percentage of Journal Citations Pre-1996

Table 2 gives the average and median citation year by year of the dissertation, and the average age of the citation. For each of the dissertation years, the average year of citation is beyond the standard online offerings of most publishers. The average citation age is above that reported in the studies discussed.

Dissertation Year	Average Citation Year	Median Citation Year	Average Age of Citation, Years
2007	1992	1996	15
2006	1994	1997	12
2005	1987	1992	18
2004	1987	1992	17
2003	1980	1986	23

Table 2. Journal citation averages by dissertation year.

Figure 2 shows the distribution of citations to journals dating before 1996, representing only those citations pre-dating the standard access provided with current online subscriptions. Average citations per dissertation are twenty or more for the periods 1990-1995 and 1980-1989, and ten for 1970-1979.

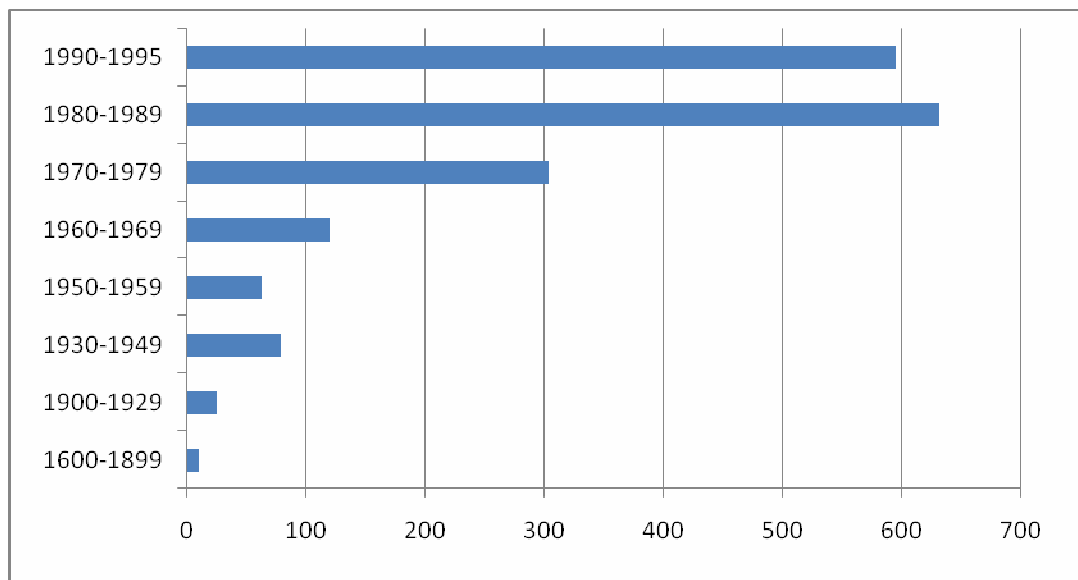


Figure 2 Journal citations by selected time periods, Total.

The two departments, Plant Biology and Zoology, did not demonstrate significant differences. For Plant Biology, the percent of citations to pre-1996 journal articles was 57%. For Zoology, the figure was 54%. All other metrics were similarly close between the two, with Plant Biology showing a slightly older average and median citation for all formats. For journals, the median citation age was identical, 1994.

Department	Average Citation Year	Median Citation Year	Median Citation Year to Journals	Average Age of Citation, Years	Percent of Journal Citations pre-1996
Plant Biology	1987	1993	1994	18	57%
Zoology	1989	1994	1994	16	54%

Table 3. Comparison of all citations, Plant Biology and Zoology

Discussion

This study demonstrates that the conventional wisdom that the sciences rely disproportionately on current sources of research is not accurate, at least for the disciplines of plant biology and zoology. The average age of the citations and the total number of citations before 1996 clearly indicate a need for access to journals beyond the typical online offerings of most publishers, for a standard subscription. Data from Table 1 do indicate that the percentage to pre-1996 journals did decline from 2003 to 2007. This could be due to the increasing online availability of recent years of journals issues, and researchers' preference for online format. Even so the high percentages of pre-1996 journal use indicate a clear need for continued availability of legacy content. Data from Figure 2 will help determine the extent of the backfiles that should be provided. Given that online access tends to increase use, availability of online journal backfiles may safely be assumed to lead to even higher citation age averages and could reverse the decreasing trend described above.

Most commercial journal publishers and university presses offer backfile purchases solely as packaged subject collections. Some publishers offer one-time purchase of these backfiles, while others also provide an annual subscription option. To accurately assess the value of subject backfile collections, citation studies of other departments will be necessary. Collecting journal title data from citation studies for all disciplines would also aid in backfile evaluation, and likely further indicate the need to pressure publishers to unbundle their electronic backfile collections and offer title-by-title selection.

An alternative is to maintain the print backfiles. This offers the advantage of title-by-title collection development. In addition, for some titles online backfile access is not available at all, so maintenance of the print copy may be the only option.

Implications are manifold. Purchasing or leasing of online science journal backfiles should not be considered solely as a desperate move to use unexpected end of year funds; in fact, it is a sound, justifiable collection development decision. Further, understanding and specifying post-cancellation rights in license agreements is crucial; purchase and local loading may be the only true guarantee an institution can rely on for the future. Similarly, retaining print back volumes, in lieu of online access, may well be worth the shelving space and maintenance costs. Lastly, support of project proposals to digitize older print volumes can be justified.

Conclusion

Citation studies have the benefit of indicating not just that a given resource was used, but *how* it was used. This is a distinct advantage over shelving counts or download reports. Citation studies of institutional authors provide a clear picture of local demand, as opposed to citation studies of a particular journal or set of journals. Citation studies of institutional dissertations provide these benefits and serve as the best indicator of the research needs of the heaviest library users.

This study shows that some science disciplines do rely on research older than typically assumed, and that in fact, in plant biology and zoology, research journals retain their value for decades. The data indicate that providing access to electronic backfiles, and/or maintaining print back volumes, in these disciplines is necessary.

This study is reproducible elsewhere and all of these significant collection development conclusions can be reasonably quickly justified by analyzing the evidence waiting to be revealed in local dissertations.

Note: The authors wish to thank Mickey Soltys, instructional support services programmer, for technical support for this study.

References

Ackerson, L.G. 2001. Is age an appropriate criterion for moving journals to storage? *Collection Management* 26 (3):63-76.

Brazzeal, B., and Fowler, R. 2005. Patterns of information use in graduate research in forestry: a citation analysis of Master's theses at Mississippi State University. *Science & Technology Libraries* 26 (2):91-106.

Delendick, T.J. 1990. Citation analysis of the literature of systematic botany: a preliminary survey. *Journal of the American Society for Information Science* 41 (7):535-543.

Edwards, S. 1999. Citation analysis as a collection development tool: a bibliometric study of polymer science theses and dissertations. *Serials Review* 25 (1):11-20.

Gooden, A.M. 2001. Citation analysis of chemistry doctoral dissertations: an Ohio State University case study. *Issues in Science and Technology Librarianship* 32. [Online]. Available: <http://www.istl.org/01-fall/refereed.html> [August 5, 2008].

Kushkowski, J.D., Parsons, K.A., and Wiese, W.H. 2003. Master's and Doctoral thesis citations: analysis and trends of a longitudinal study. *portal: Libraries and the Academy*, 3 (3):459-479.

Musser, L.R., and Conkling, T.W. 1996. Characteristics of engineering citations. *Science and Technology Libraries* 15 (4):41-49.

Salisbury, L., Bajwa, V., and Dillon, S.L. 2007. University of Arkansas Food Science Faculty publications and the characteristics of their cited references: a bibliometric study. *Journal of Agricultural and Food Information* 8 (4):21-33.

Vallmitjana, N., and Sabaté, L.G. 2008. Citation analysis of Ph.D. dissertation references as a tool for collection management in an academic chemistry library. *College and Research Libraries* 69 (1):72-81.

Walcott, R. 1994. Local citation studies - a shortcut to local knowledge. *Science and Technology Libraries* 14 (3):1-14.

Williams, V.K., and Fletcher, C.L. 2006. Materials used by Master's students in engineering and implications for collection development: a citation analysis. *Issues in Science and Technology Librarianship* 45. [Online]. Available: <http://www.istl.org/06-winter/refereed1.html> [August 5, 2008].