Evidence Based Library and Information Practice

Article

Assessing the Fitness of an Academic Library for Doctoral Research

Susan Edwards
Head, Psychology, Education and Social Welfare Libraries
University of California, Berkeley
Berkeley, California, United States of America
Email: sedwards@library.berkeley.edu

Lynn Jones
Reference Coordinator
Doe Library
University of California, Berkeley
Berkeley, California, United States of America
Email: ljones@library.berkeley.edu

Abstract

Objective – At the University of California, Berkeley (UCB), researchers compared how well the library collections supported doctoral research in the three related disciplines of education, psychology, and social welfare. The goal of this project was to gather empirical data to answer questions about materials cited in dissertations, including ownership, age of materials and disciplinary usage.

Methods – Researchers analyzed the bibliographies of doctoral dissertations from three academic departments at UCB: education (2009-2010), psychology (2009-2010), and social welfare (2009-2011). The sampling methodology used a systematic sample with a random start. To achieve a 95% (+/-4%) confidence interval, the sample included a total of 3,372 citations from 107 dissertations. Researchers consulted with a statistician to determine the statistical significance of the results. The test for the age of citation used a signed

ranks test, which is typical for ordinal data or skewed interval data. The test for ownership was a chi-square test, which is typical for nominal data or dichotomous data.

Results – Researchers determined that a very high percentage of the cited journals were owned or licensed by the Library. The ownership rate for cited journals was 97% for both education and social welfare, and 99% for psychology. There was a statistically significant difference between the three disciplines, with psychology better supported than either education (p=.02) or social welfare (p=.01). However, since ownership rates for journals in all three disciplines were extremely high, this was not a meaningful difference. For books, the researchers found a significantly smaller percentage of books owned in social welfare compared to either education (p=.00) or psychology (p=.00). We found no significant difference between the percentages of books owned in psychology versus education (p=.27). Psychology students cited the highest percentage of journals while education students cited the highest percentage of books. Psychology students cited almost no free web resources, but education and social welfare students did cite free web resources (primarily government documents, working papers, or non-governmental organization reports). All three disciplines cited older material than anticipated.

Conclusions – The citation analysis, while time-consuming, provided new and important information about the use of the Library's collections and the level of support the collections afford doctoral students in the three related disciplines of education, psychology and social welfare. This data has informed collections-related decisions including format purchases and fund allocations.

Introduction

The reputation of research libraries and their collections have long been entwined, with the size of the latter often held as a key indicator of the quality of the former. However, traditional measures of collection strength -- dollars spent, number of volumes added, percent of scholarly publishing acquired, range of languages collected, and even circulation, interlibrary loan requests, and usage data-- are no longer enough. It is not that, as the *Star Wars* character Yoda says, "size matters not," (Kurtz & Kershner, 1980) but that shrinking buying power and a growing need to demonstrate return on investment requires we understand how well the

library supports the research of its students and faculty.

University of California – Berkeley's (UCB) state-funded collections budget has not increased since 2001. Inflation, the decreasing value of the United States dollar abroad, the rising cost of journals, and the need to handle new formats such as data have increased pressure on collections budgets. This challenging budget situation has made it more important than ever to understand how well the collections support the disciplines, and whether some disciplines, or some format types, have been disproportionately impacted.

Citation analysis of dissertations offers a reliable source of data to assess how well the collections meet the needs of

doctoral students, a core user group of the library. Citing a source in a dissertation indicates that the source was of value to the student; this is a conclusion that cannot be made from usage statistics alone. Since education, psychology, and social welfare all offer doctoral programs at Berkeley, data derived from dissertation citations allows for meaningful comparisons between these disciplines. A large literature supports the notion that doctoral student research, in addition to being valuable in its own right, serves as a reasonable proxy for faculty research patterns (Zipp, 1996).

With the current emphasis on resource access over resource ownership, citation analysis of ownership and licenses may seem anachronistic. The authors are well aware of the importance of access, and understand that no library can meet all the needs of doctoral level researchers. However, decisions must be made about how to allocate funds, both between and within disciplines, and the researchers wanted scholarly practice at the home institution to inform collections assessment and analysis.

Aims

The goal of this project was to gather empirical data to answer the following questions:

1. What percentage of cited material is owned or licensed by the library? Does it vary by format and/or discipline?
2. How does Berkeley compare to peer research institutions in its level of support for graduate research in these disciplines?
3. What is the ratio of books to journals cited, by discipline?
4. Is there an emerging pattern of free web resources being cited, by discipline?
5. How old is the material being cited? Does it vary by discipline or format?
6. What are the most frequently cited journals overall, and by discipline?

Literature Review

A substantial literature exists on the topic of citation analysis, although not all publications in that pool were relevant to our research due to differences in methodology, disciplinary areas, and assumptions. Sources used to inform this study primarily focus on dissertation citation analysis in research-level library collections in the social sciences, in particular in the fields of psychology, education, and social welfare. Hoffmann & Doucette (2012) extract the methodological details of 34 citation analysis studies and compare them. Among the variables most commonly analyzed by these studies are: percent of holdings owned or licensed, citation age, frequency of journals cited, citation by title, and type of document cited. The article is quite valuable for those designing a citation analysis study.

The earliest relevant citation analysis research found was the seminal work by Peritz and Sor (1990). Their original findings supporting the use of dissertation citations as a basis for collection decision-making have been replicated and substantiated by many others over the past 23 years. Zipp's (1996) research reinforces the claim that the research of graduate students and faculty are sufficiently similar that analysis of dissertations can serve as an

adequate proxy for analyzing the bibliographies of faculty publications.

A common goal of citation analysis studies is to identify core journals for selection (or deselection) purposes. Waugh and Ruppel (2004) summarize reasons why citation analysis of dissertations is one of the most reliable ways to determine journal usage. They also make the point that a list of most heavily used journals can be helpful to a department in several other ways, such as suggesting places for authors to publish, to help weigh the value of a publication in the academic review process, and to help decide which library fund should pay for which journals. Wirth and Mellinger (2012) present innovative ideas for using citation analysis data to improve online subject guides. Thomas's (2000) research on the citation behaviour of 1,024 social work master's students is typical in its research questions, including which journals were cited, and the median age of citations, which new titles should be purchased, and which funds were used to purchase cited materials.

An ongoing controversy in the use of citation analysis of student work is the so-called 'convenience bias'. This is the argument that students will use journals available locally rather than using 'the best' articles which might require going into the stacks or even using interlibrary borrowing (Sexton, 2006). This argument was not a concern for the present project for two reasons: 1) the researchers assume that it would be equally true (or not) across the three disciplines we studied, therefore making the comparative analysis still meaningful, and 2) the researchers want to know what types of resources students are using (e.g., the mix of books to journals by discipline, or the

age of material cited), even if these are the most convenient materials.

Some question the premise that citation analysis of dissertation references is an appropriate basis for collection building (Haycock, 2004). Beile, Boote, and Killingsworth (2004) ask, "is it reasonable to conclude… that research collections that contain the majority of cited items are sufficient for doctoral level research?" (p. 348). On the contrary, they claim that basing collection decisions on usage by a single institution's students will lead to a "skewed list of journals" because "doctoral students simply do not possess sufficient knowledge of information resources, expertise in mining the literature of the field, or the ability to… create quality bibliographies" (Beile, Boote & Killingsworth, 2003, p. 12). The researchers agree that there are limitations to the value of citation analysis for deciding which journals to add or cancel, but were not using this data to make individual title decisions.

Based on the literature review for this study, there appear to be differences between the citation behaviour of masters' level students and doctoral candidates. For example, when comparing the citation behaviour of faculty and masters level students in biology, Pancheshnikov (2007) found that the journals cited by each group are quite different from one another and that faculty cite a broader array of sources than master's students. Pancheshnikov concluded that faculty work should be the basis of collection decisions. As noted above, doctoral research more closely matches that of faculty. Thus, the current study focuses only on doctoral student dissertations, not a wider range of graduate student

work. While there may be disagreement over how well student dissertation research mimics that of faculty, it is the researchers' contention in this article that the doctoral research of graduate students is valuable in its own right and must be supported by research library collections.

Methods

The researchers analyzed the bibliographies of all doctoral dissertations from UCB from three disciplines: education (2009-2010), psychology (2009-2010), and social welfare (2009-2011). Social welfare has a significantly smaller number of doctoral students (see Table 1) so an additional year of data were used to make the sample sizes more comparable. In this study we consulted an evaluator to determine the sampling methodology, and determined that a systematic sample with a random start would be appropriate. To achieve a 95% (+/-4%) confidence interval the researchers sampled a total of 3,372 citations from 107 dissertations. We also consulted a statistician to determine statistical significance of the results. The test for the age of citation used a signed ranks test, which is typical for ordinal data or skewed interval data. The test for ownership was a chi-square test, which is typical for nominal data or dichotomous data.

The researchers gathered lists of students and their dissertations from each department's administrative office. Most dissertations were located in the library catalogue or Digital Dissertations database (http://search.proquest.com/dissertations/). Eleven dissertations were unavailable and therefore were not sampled, possibly introducing a sampling bias for which we have not corrected. As of 2009 all of Berkeley's dissertations are available in electronic format only, and students can choose to make them available immediately, or to embargo them for two years or more with justification.

For each discipline the researchers selected a random number from 1 to 5. The citation corresponding to that number, and every fifth citation after, was selected from each bibliography and entered into an Excel spreadsheet created for that discipline. We used numbers in place of authors' names in project spreadsheets to identify dissertations in order to preserve the anonymity of graduate students. Every sampled citation was listed under the appropriate assigned student number; each citation was also assigned a unique identifying number to allow us to sort and resort the data and still be able to restore the original sequence as needed.

Table 1
Number of Citations in Doctoral Dissertations by Discipline

Discipline	Total dissertations sampled	Number of citations in final sample
Education	57	1,340
Psychology	28	644
Social Welfare	22	1,388
Totals	**107**	**3,372**

Student library employees conducted preliminary searching of citations in Berkeley's online catalogue and subscription databases licensed locally, as well as in resources of the Library's consortial partners to which Berkeley's students have access. They entered ownership or licensing status into the spreadsheets, along with the information for titles which were freely available on the web, or for those not owned or licensed. Citations that students were unable to verify were later researched by the authors and added to the spreadsheets.

Researchers then analyzed the collected data to determine the following:

- the percentage of cited journals and monographs owned or licensed by UCB Library
- median number of citations per dissertation in each department
- the percentage of books vs. journals cited, sorted by department
- journal titles ranked by frequency cited, by department
- the top ten most-cited journals, by department, annotated with Ulrich's subject designation
- the median age of citations (grouped by journals or books) in dissertations in each department. This was calculated as the difference between the year of submission of the dissertation and year of publication of the item cited.

Results

Our research confirmed that the journals owned or currently licensed by UCB Library provided a very high percentage of articles cited by students in all three disciplines. As shown in table 2, the ownership rate for cited journals was 97% for both education and social welfare, and 99% for psychology. There was a statistically significant difference between the three disciplines, meaning it was unlikely to have occurred by chance in the samples, with psychology better supported than either education (p=.02) or social welfare (p=.01). However, since ownership rates for journals in all three disciplines were extremely high, this was not a meaningful difference.

For books, the researchers found a significantly smaller percentage of books owned in social welfare compared to either education (p=.00) or psychology (p=.00) but found no significant difference between the percentages of books owned in psychology versus education (p=.27).

The mix of journals to books varied significantly by discipline, with psychology dissertations citing the highest percentage of journals, and education dissertations citing the highest percentage of books. As shown in table 3, psychology students cited almost no free web resources, but education and social welfare students did cite free online resources with some frequency, approximately 7% of their total citations. The free web resources used were primarily government documents or reports from non-governmental or advocacy organizations, while these types of materials were rarely cited by students in psychology students, who overwhelmingly used peer-reviewed articles.

One surprising finding was the age of material students cited, shown in table 4. Not only did all three disciplines cite material that was older than anticipated, the assumption that psychology would rely almost exclusively on recent research was not supported. On the

Table 2

Journals and Books Owned or Licensed by the Library as Cited in Doctoral Dissertations

Discipline	Journals	Books
Education	97%	86%
Psychology	99%	87%
Social Welfare	97%	72%

Table 3

Type of Sources Cited in Doctoral Dissertations

Discipline	Journals	Books	Web sources [government documents, etc.]
Education	46%	47%	7%
Psychology	84%	15%	<1%
Social Welfare	59%	33%	8%

Table 4

Median age of citations in doctoral dissertations

Discipline	Journals	Books	Combined
Education	11 years	13 years	11 years
Psychology	8 years	14 years	9 years
Social Welfare	10 years	11 years	10 years

contrary, half the sources cited by psychology doctoral students were more than nine years old. In general, education students cited older journals than psychology, but we found no statistically significant difference for books between these two disciplines. Social welfare and education students cited older journals than psychology, but surprisingly we found that psychology students cited older books than social welfare. The difference in citation age between education and social welfare was not statistically significant for books or journals.

One of the most surprising findings was the strong cross-disciplinary nature of the research in these three disciplines. As shown in table 5, social welfare and education doctoral students frequently cited journals in psychology instead of the high impact journals in their respective fields, while psychology doctoral students relied on journals classified as psychiatry and neurology, and sciences: comprehensive works

according to the Ulrichsweb Global Serials Directory (http://ulrichsweb.serialssolutions.com/) subject classifications. The neuroscience orientation of UCB's Psychology Department was well known to the selector for that area, but the heavy reliance on psychology journals by education and social welfare students was a surprise both to the selector and the faculty who serve as doctoral advisers. This finding has implications for fund allocations and for the importance of collaborative review in serials cancellations projects since the discipline which funds the title may be different from the one(s) that are using it.

Discussion

As part of a major research institution, the UCB Library has a tradition of outstanding collections and a mandate to support doctoral level research. In an

era of rising serial costs and deepening fiscal
constraints, it is imperative to analyze how well
the collections still support doctoral student
research and whether departments were
equitably supported. The results show that UCB
Library met virtually all of the journal demand

Table 5
Most Frequently Cited Journals in Doctoral Dissertations by Discipline

Rank	Education	Psychology	Social Welfare
1st	*Journal of Personality and Social Psychology*	*Neuroimage*	*Child Development*
2nd	*Child Development*	Journal of Personality and Social Psychology	*Developmental Psychology*
3rd	Journal of Educational Psychology	*Journal of Neuroscience*	Children and Youth Services Review
4th	*Developmental Psychology*	*Nature* *Neuropsychologia* *Nature Neuroscience*	*American Psychologist* *Development and Psychopathology*
5th	Journal of Research in Science Teaching Journal of the Learning Sciences Reading Research Quarterly	*Neuron* *Science*	Future of Children Child Abuse & Neglect
6th	American Educational Research Journal Applied Psychological Measurement Educational Psychologist	*Journal of Neurophysiology* Journal of Cognitive Neuroscience	Child Welfare *American Sociological Review* *Journal of Consulting & Clinical Psychology* *Pediatrics* *Social Science and Medicine*

Note: Italicized titles are from outside the discipline.

in psychology, education, and social welfare, the three disciplines being studies. However, the Library does not provide as high a percentage of monographs as it does journals. In fact, the data show weaker support for monograph budgets for all three disciplines, which leads the researchers to ask an important question: has the Library sacrificed monographs to maintain journal coverage, and is this wise?

One of the researchers' initial goals was to learn how well the UCB Library supports the research of doctoral students and how well that compares with library support provided by its peer institutions. The literature review did not reveal data from peer research libraries in these disciplines with a comparable methodology, so we were unable to make this comparison. Nevertheless, this data is important for internal benchmarking within UCB Library, as one measure for comparing for how well the library supports the research of different departments. We note that 86-87% of monographs used by education and psychology doctoral students are available at UCB Library while only 72% of monographs for social welfare are included in the UCB Library collection. This discrepancy is prompting meaningful discussions about budgetary equity between disciplines.

With this data, the researchers hope to begin a conversation within our own institution and with peer libraries about how much support of research is enough. What is an acceptable level of doctoral research support? Two other findings were significant for collection development decisions. First, graduate students in all three disciplines cited material significantly older than

expected. The median age of citations in education and social welfare was about ten years, meaning a full 50% of citations were to material more than ten years old. For psychology, the median age of citations was nine years old. This finding contradicted the conventional wisdom that sciences such as psychology do not cite older materials. With this new information, UCB Library's selectors in these disciplines intend to purchase more online back files, and to consider more carefully our deselection decisions for older material.

Citations from these three disciplines showed more usage of psychology journals by education and social welfare than we had assumed. Three of the top four journals cited by education students and four of the top five journals cited by social welfare students are classified in psychology by Ulrich's. Decisions to cancel journals in psychology based on the usage behaviour only of psychology graduate students would hamper the research of students in education, social welfare, and perhaps many other departments. This finding demonstrates the need for data from multiple departments and multiple sources in making the best collection development decisions, and has implications for disciplinary budget allocations as well.

Next Steps

This study produced enough actionable information that the citation analysis will be continued for dissertations in business, economics, political science, and history. This next phase received a research grant to hire students for data entry and a library assistant for bibliographic verification. Using the same project design and statistical

methodology will enable a comparison between all seven departments.

A careful analysis of the citation formatting in the bibliographies provides a fertile topic for outreach to departments, and possible workshops for graduate students. Citation errors were common, in particular how to cite edited works and primary sources. Understanding citation errors— both the rate and types of citation errors, which were not tracked in this study— by department or advisor would help librarians design appropriate and targeted instruction.

Four immediate uses of this data are planned. The UCB Library has begun to purchase additional journal back files in psychology, given the strong reliance on journals versus monographs in that department. Some monographic funds are being re-allocated from psychology to social welfare to correct the imbalance found. Titles cited but not owned or licensed by UCB Library are being evaluated to determine whether they should be acquired, and we are working closely with faculty and students to better understand what types of monographs they need and what areas of the collection need strengthening.

In the next iteration of this research, more detailed information on types of materials cited will be collected, not just monographs, journals, and websites, but also dissertations, news sources, archives, data sources and other genres. Tracking changes in what material types are being cited by graduate students

provides insight into changing trends in scholarship and local research practice. This project did not assess how well doctoral student citation behaviour matches that of faculty. Since there is a controversy in the literature over this point, it would be useful to investigate this correlation for the design of future citation research at UCB Library.

Conclusion

This pilot project provided new and important information about the use of collections by doctoral students in education, psychology and social welfare UCB. The researchers learned that our journal coverage is excellent in all three disciplines, but that our monographic coverage for all three disciplines is less complete, particularly in social welfare. Cross-disciplinary use of journals is greater than expected, and all three disciplines use older materials than assumed. This data will be used to inform budget reallocation, collection development and management.

In addition, designing and conducting a quantitative social sciences research project has made us much more sensitive to the challenges and complexities that face students. Our research project has showed us that the literature review, which librarians primarily focus on, is just one element of a much larger research process. Becoming deeply engaged with our own research has equipped us to have different and more sophisticated conversations with students, fellow librarians, faculty, and academic deans and chairs about academic research.

References

Beile, P. M., Boote, D. N., & Killingsworth, E. K. (2003). Characteristics of education doctoral dissertation references: An inter-institutional analysis of review of literature citations. Presented at the American Education Research Association Annual Conference, Chicago, IL: ERIC.

Beile, P. M., Boote, D. N., & Killingsworth, E. K. (2004). A microscope or a mirror?: A question of study validity regarding the use of dissertation citation analysis for evaluating research collections. *Journal of Academic Librarianship*, *30*(5), 347-353. http://dx.doi.org/10.1016/j.acalib .2004.06.001

Haycock, L. A. (2004). Citation analysis of education dissertations for collection development. *Library Resources and Technical Services*, *48*(2), 102-106. http://alcts.metapress.com/conte nt/K47N44/

Hoffmann, K., & Doucette, L. (2012). A review of citation analysis methodologies for collection management. *College & Research Libraries*, *73*(4), 321-335. http://crl.acrl.org/

Kurtz, G. (Producer), & Kershner, I. (Director). (1980). *Star wars V: The empire strikes back* [Motion picture]. United States: Lucasfilm.

Pancheshnikov, Y. (2007). A comparison of literature citations in faculty publications and student theses as indicators of collection use and a background for collection management at a university library. *Journal of Academic Librarianship*, *33*(6), 674–683. http://dx.doi.org/10.1016/j.acalib .2007.09.011

Peritz, B. C., & Sor, D. (1990). The use of libraries by graduate students in psychology as indicated by citations. *Collection Management*, *12*(3/4), 11-23. http://dx.doi.org/10.1300/J105v1 2n03_02

Sexton, E. (2006). Journal use by graduate students as indicated by master's theses bibliographies at an urban commuter college, 1991-2004. *Behavioral & Social Sciences Librarian*, *24*(2), 93-111. http://dx.doi.org/10.1300/J103v2 4n02_04

Thomas, J. (2000). Never enough. *Behavioral & Social Sciences Librarian*, *19*(1), 1–16. http://dx.doi.org/10.1300/J103v1 9n01_01

Vallmitjana, N., & Sabaté, L. G. (2008). Citation analysis of Ph.D. dissertation references as a tool for collection management in an academic chemistry library. *College & Research Libraries*, *69*(1), 72 -82. http://crl.acrl.org/

Waugh, C. K., & Ruppel, M. (2004). Citation analysis of dissertation, thesis, and research paper references in workforce education and development.

Journal of Academic Librarianship, *30*(4), 276-284. http://dx.doi.org/10.1016/j.acalib.2004.04 .003

Zipp, L. (1996). Thesis and dissertation citations as indicators of faculty research use of university library journal collections. *Library Resources and Technical Services*, *40*(4), 335-342. http://alcts.metapress.com/content/K47 N44/