

Citation Analysis of Masters Theses: as a Tool for Collection Development in Academic Libraries

Gunasekera, Chamani¹

Abstract

This study analyses the citations of Masters Theses on Sociology at the University of Peradeniya submitted from 1995 to 2012 to determine the format and age of materials used and most frequently cited journals. The citation analysis method was used as the data collection method and the citations were extracted from the title pages and reference lists of each of the theses. Data obtained from 12 masters theses were examined in June, 2013. The study found that 56% of cited items were monographs followed by 22% were journals, 8.5% were reports, 4% were web resources and 3.7% were conference proceedings. This is contrary with other citation analysis, which found that journals are the most frequently used format. The study also revealed that nine journal titles are the most frequently cited journals by sociology graduates. The study indicated that the average age of materials used was 10- 20 years. The findings and implications of the research for collection development have been discussed. This study could serve as a collection development tool that can be used as a model for the library to identify the primary sources for acquisitions and also as a guide for collection maintenance.

Keywords: Citation analysis, Collection development tool, References analysis, Academic libraries

Introduction

Citation analysis is a branch of bibliometrics that examines the citations found in publications such as journal articles and books (Georgas and Cullars, 2005). Common approaches in the literature are to analyze faculty publications or graduate student theses for a particular department at a particular institution. Items cited in students' dissertations show the resources they used in their dissertation writing. This type of analysis can yield valuable insights into local citing trends, such as types and number of documents cited, age of citations, and changes in citation patterns over time (Ashman, 2009).

¹ Senior Assistant Librarian, University of Peradeniya, Sri Lanka. Email: chamanig@yahoo.com

Academic libraries particularly the university libraries are constantly under pressure to manage funds effectively and in a consistent manner with teaching and research needs. Libraries that want to provide the greatest possible access to the greatest possible breadth of resources need to know and understand the resource needs of their users. Faculty research interests may be identified from lists of faculty publications, faculty handbooks, and faculty requests for material purchase. But research needs of graduate students can be more difficult to understand. Analysis of dissertation and thesis reference lists is one approach used to measure library use by graduate students, who are traditionally frequent and heavy library users. In general dissertations are considered as “valuable road signs” to the literature of a discipline (Cox, 2008). Dissertations clearly indicate the needs of graduate students and also indicate the research specialties of the faculty and department as a whole.

Citation analysis is extensively used as a tool in collection development. With ever increasing cost of library resources and the limited library budgets, the librarians are faced with the challenge of make decisions on collection development. This is increasingly difficult as journal subscriptions is rising up continuously when get limited shrinking budget. Citation analysis can act as a tool for selecting and de selecting materials as it provides insight into the materials that are selected by various user groups. It can be a useful technique for identifying potential collection development weaknesses also (Rethiefsen, 2007). It further helps for the librarian to identify unique information that was not found from other library statistics.

Citation analysis is based on the assumption that documents cited by a researcher have been used in the research process. Consequently, this approach helps clarify both the information needs of researchers and what should be contained in a research library collection. Major use of citation analysis pertinent to collection evaluation include identify the core collection, using citation as a check list, ranking journals and analyzing a disciplines structure to assist collection development decision making (Waugh,2004) .

Many libraries are under pressure to make difficult decisions regarding older materials sitting on shelves, including print journals. Lacking of shelving space is one of the critical problems faced by library University of Peradeniya at present. In order to make decisions on what materials should be in the shelves and what can be moved out of the shelves can be assessed by examining the age of cited resources by library users.

Present study analyzed citations in the Sociology Masters theses submitted to the Main Library, University of Peradeniya during the period of 1995 - 2012. The purpose of this research is to analyze citation pattern of sociology theses and built up indicators to support library's collection development. Specifically, it intends to:

1. Identify the different types of resources cited by the masters' students;
2. Ascertain the age of the resources cited in relation to the year theses were published; and
3. Determine the most frequently cited journals within sociology theses

In many academic libraries, funds for purchases are allocated by discipline, and within the discipline the allocation is divided between serial and monograph. The ratio of serials to monograph use was explored in comparison to the allocation of the library budget. So, awareness of graduate citation patterns may assist in developing such formulas in the collection development.

Examining the age of cited references will help to develop storage and retention policies of the library. Analysis of citation data can be used to create core journal lists, which can be used to make collection decisions in support of research and teaching. Citation data may be used in developing core title lists of journals to maintain when flat or shrinking budget allocations require journal cancellations.

Citation Analysis and Collection Development

With a “collection management” citation analysis, the underlying premise is that “the more frequently cited publications are the more valuable, will continue to be used heavily, and, consequently, are more important to have in the library collections” (Johnson, 2007). Collection-focused activities that make use of citation data include examining print versus electronic usage to determine the impact of electronic journals, compiling lists of most- and least cited journals and local holdings to make acquisition or cancellation decisions, and examining age of cited references to help develop storage and retention policies (Smith, 2003).

Gross and Gross (1927) used for the first time this citation analysis as a tool for identifying the core journals in a subject based on counting the citations given at the end of each article from a group of a primary journals. Afterwards, a number of such studies were carried out in citations in dissertations/theses, books, primary journals and reviewing journals.

Buchanan and Herubel (1994) and Edwards (1999) assert that the ratio of journal articles to monographs is lower in the humanities and social sciences. Their study of citations in political science dissertations supports this view of the social sciences (46.7%) of citations to monographs, but neither study cites research to support this contention with regard to other social sciences or the humanities.

Citation analyses have been used as a collection development tool within the disciplines of engineering, linguistics, education, sciences and psychology (Williams & Fletcher, 2006; Georgas & Cullars, 2005; Haycock, 2004; LaBonte, 2005; Sylvia, 1998). Multi-disciplinary studies by Kushkowsky, Parsons and Wiese (2003) and E.T.Smith (2003), and a survey of English literature masters theses by Slutz (1997) indicate that selected humanities and social sciences are dependent on monograph. Slutz calculated that 50.2% of citations in English Masters theses were to books and 27.4% to journal articles. However, all three studies examined comparatively small data sets for humanities and social sciences. For

example, their data sets for humanities were n=approximately 449; n=377; and n=938 citations respectively.

Several studies have suggested that within this general pattern, the age distributions for humanities and social sciences are flatter because of the greater use of older materials. One method of comparison is to calculate the mean and median ages, which – according to this hypothesis – would be higher in humanities and social sciences. Kushkowski, Parsons and Wiese (2003) report an overall mean of 12.3 years for their multidisciplinary survey and 18.1 years for humanities. They also report an overall median of eight years and ten years for humanities, but the latter statistic is also reported for some science disciplines. This prevents them from distinguishing between an emphasis in the natural sciences on recently published research and an emphasis in the humanities and social sciences on older “foundational” publications.

Journals are one of the most commonly used research tools (Pancheshnikov, 2007; Rethlefsen, 2007; Williams & Fletcher, 2006). Citation analysis can be used to rank, evaluate and categorize journals based on their frequency of article citation. It can help identify areas of weakness within a collection (Rethlefsen, 2007).

Citation analysis is an effective tool but it has limitations. MacRoberts and MacRoberts (1989) stated that 'most authors simply did not cite the majority of their influences, and none cited all influences'. Graduate students may include references even if a work is marginal or irrelevant to their theses (Sylvia, 1998). In turn, students may not cite important research if it is not published in their native language or cannot be found locally (MacRoberts & MacRoberts, 1989, Sylvia, 1998). Budd (1986) asserted the use of citation analysis in the practice of collection management in the humanities is still something of a question marked and he concluded that references to books are very diffuse. The most frequently cited primary monograph and secondary monograph are identified and compared with standard lists and other citation studies. At the end he stressed that “small clusters of frequently cited materials were identified, but that the citation analysis was

less successful in dealing with “fringe” areas. Therefore researchers must use caution when drawing conclusions and making recommendations based solely on citation analysis data. Broadus (1985) recommended that using citation analysis in addition to considering in-house use, circulation, and interlibrary loans to determine the usefulness of items.

Methodology

In this study Masters Theses of Sociology submitted during the period of 1995-2012 to the Main Library, University of Peradeniya were examined. Citations were extracted from the title pages and reference lists of each of the thesis were examined. Each thesis’s title page and reference section were photocopied and the data was feed into the SPSS package for analyzing. Data extracted included the year of the materials were submitted, title of the thesis, author, and year of acceptance, format of material cited, date of publication and name of cited journal.

The material cited were categories as; journals, books, conference proceedings, reports including working papers, research paper, annual reports, and technical reports, theses and dissertations, web resources, government publications and miscellaneous (which includes year books, newspapers, manual, bulletins, guide, newsletters, magazine ,personal communications, unpublished materials, and other). Data obtained from twelve (12) masters theses were examined in June, 2013 .The data extracted were analyzed using descriptive statistics which includes frequency and percentage presented in tables and figures.

Findings and Discussion

The results were presented based on the objectives of the study which include the format of resources the students cited, the age of citation with reference to the year of the theses were published and the most cited journals by the students. From twelve (12) masters theses one thousand two hundred and thirty two (1232) references were found. In this process, every reference made at one time has been counted as one citation.

Format used

Materials were grouped into eight format categories for analysis: journals, books, conference proceedings, reports including working papers, research paper, annual reports, and technical reports, theses and dissertations, web resources, government publications and miscellaneous (which includes year books, newspapers, manual, bulletins, guide, newsletters, magazine, personal communications, unpublished materials, and other).

Following are some of the key findings discovered in relation to the format of cited materials.

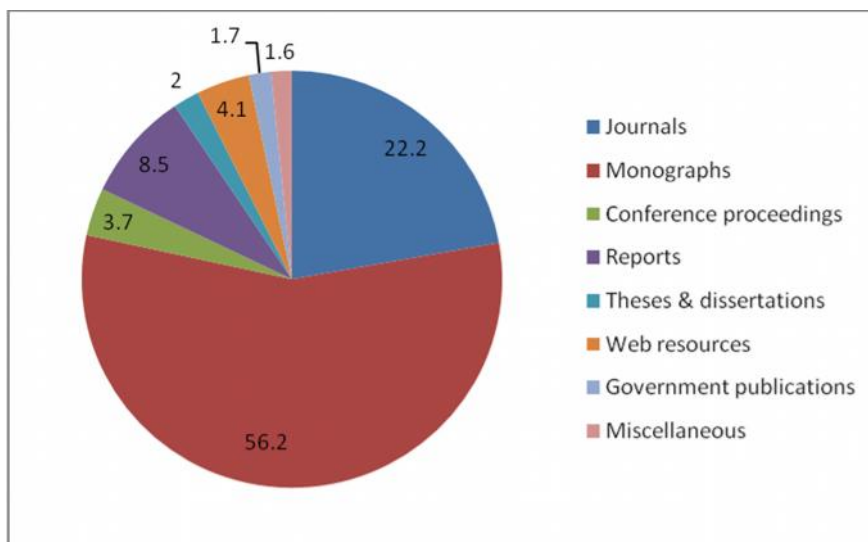


Figure 1: Types of resources cited

Table 1: Citation According to Types of Information Materials

S/N	Material type	No. of citations	Percentage (%)	Cumulative percentage
1	Journals	274	22.2	22.2
2	Monographs	692	56.2	78.4
3	Conference proceedings	45	3.7	82.1
4	Reports	105	8.5	90.6
5	Theses & dissertations	25	2.0	92.6
6	Web resources	50	4.1	96.7
7	Government publications	21	1.7	98.4
8	Miscellaneous	20	1.6	100
Total		1232	100	

As indicated in the figure 1 and table 1 fifty six percent (56%) of the total citations were monograph, followed by 22% of journals, 8.5% of reports, 4% of web resources and 2% of theses and dissertations. The results of this study is contradictory from the previous studies carried out which found journals were the most commonly cited format , with other format varying widely. (Rieb 1993; Kim 2002; Kushkowski et al. 2003; Musser and Conkling 1996)

Material Age

The analysis of citations by age of cited documents reveal useful life of documents. This period of citation of the journals are popularly referred as “half life of periodicals” or often quoted as “obsolescence of use of literature” (Shafi, 2001) For calculating the citation age, the difference between the date of the citation and the date of the publication in which it was cited considered rather than simply analyzing publication date. In the current study, percentages were used to describe items to make the data more meaningful. Figure 2 shows the percentages of age of cited materials by master students on sociology.

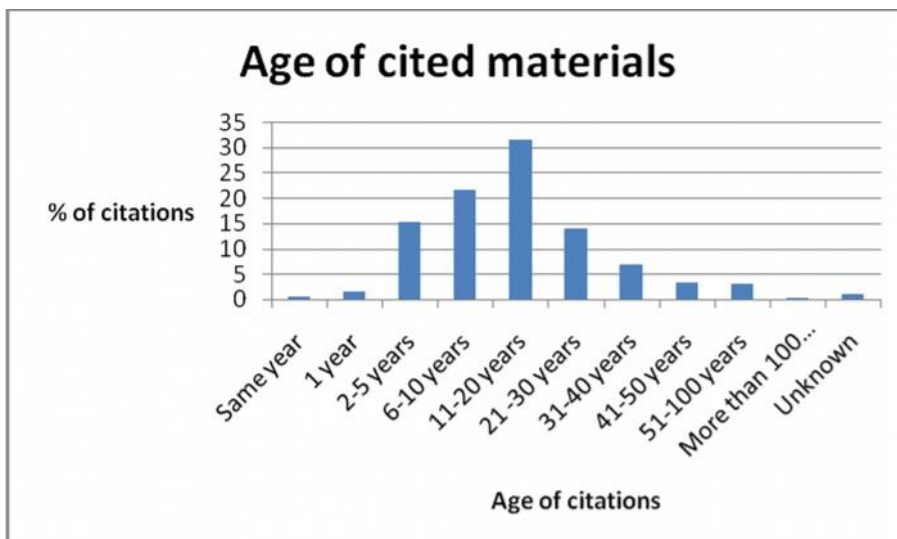


Figure 2: Age distribution of cited resources

Understanding the extent to which library users rely on older materials can be useful in determining which materials can be moved to remote storage (Ackerson 2001). Musser and Conkling (1996) and Kushkowsky et al. (2003) found that the majority of materials cited were less than eight years old.

Table 2: Age of cited materials

No. of years	No. of citations	Percentage (%)	Cumulative percentage
Same year	6	0.5	0.5
1 year	20	1.6	2.1
2-5 years	188	15.3	17.4
6-10 years	268	21.8	39.2
11-20 years	389	31.6	70.7
21 -30 years	172	14.0	84.6
31-40 years	85	6.9	91.6
41-50 years	39	3.2	94.7

51-100 years	37	3.0	97.7
More than 100 years	4	0.3	98.1
Unknown	24	1.0	100.0
Total	1232	100.0	

As indicated in the table 2 and figure 2 thirty nine percent (39%) of materials cited were published within the last ten years. It further revealed that seventy one percent (71%) of materials were published within the last 20 years and 85 % of materials within the last 30 years. The oldest item cited was a monograph which published in 1681.

Ranked List of Journals

Journals are essential for research, but due to the ever increasing demands for the journals librarians are struggling to select most relevant titles by studying the quality, usefulness and suitability to particular group of users. The ranking list of journals is a practical tool that helps librarians to select journals of maximum utility in relation to their coverage of new and important literature in a particular subject area.

In the present study, citation analysis has been used for preparing rank list of journals in the discipline of Sociology. This Sociology Master’s Thesis study contained 274 journal citations from 105 journal titles. The journals were placed into three zones according to Bradford’s Law of journal scattering. The number of journal citations was divided by 3 to determine how many citations should be included in each zone. (See table 3)

Table 3: Percentage of no. of journals and citations in each Bradford Zones

Zone	No. of Journals	No. of Citations	% of Journals	% of Citations
1	9	90	8.0	33.0
2	22	85	21.0	31.0
3	75	99	71.00	36.00
Total	106	274	100	100

According to our application of Bradford’s Law, 09 journals are considered as core journals. Journals in the first and second zones encompass together 175 citations in 31 journals. The remaining almost 71% of the journals were cited less than 2 times, and 49% of the journals were cited only once. The second zone contained 85 citations in 22 journal titles (See table 5) and the third zone had the remaining 24 journal titles with 2 citations each and 51 journal titles with one citation each. Majority of the journal titles (n:51=49%) has only one citation.

Table 4 shows the 9 journal titles comes under first zone with 90 citations.

Table 4: Journals in the first zone

S/N	Journal Name	No. of citations	Cumulative	Rank
1	Economic Review	28	28	1
2	World development	14	42	2
3	American Sociological Review	9	51	3
4	American Journal of Sociology	8	59	4
5	American Anthropologist	7	66	5
6	Modern Ceylon studies	6	72	6
7	Population & Development review	6	78	6
8	Population studies	6	84	6
9	Sri Lanka journal of Social sciences	6	90	6

Table 5: Second zone of the journals

S/N	Journal name	No.of citations	Cumulative	Rank
1	Social Science Medical	5	95	7
2	Development & Change	5	100	7
3	Asia Survey	5	105	7

4	The Journal of Asian Studies	5	110	7
5	Samaja Vimasuma	5	115	7
6	International Migration Review	5	120	7
7	American JI. of public health	4	124	8
8	Ceylon JI.of historical& social studies	4	128	8
9	Gender & society	4	132	8
10	Jl. Of international development	4	136	8
11	Journal of peasant studies	4	140	8
12	American political science review	4	144	8
13	American economic review	4	148	8
14	American Ethnologist	3	151	9
15	Asia pacific migration JI.	3	154	9
16	Ceylon medical JI.	3	157	9
17	Feminist review	3	160	9
18	JCBRAS	3	163	9
19	Jl. of contemporary Asia	3	166	9
20	Jl.of development economics	3	169	9
21	Marga	3	172	9
22	Pravada	3	175	9

Garfield (1977) suggested that rankings developed from total citation counts could be a variable in developing core journal lists. However in the present study, the rank list of journals in the field of Sociology reveals that journal citations cited by researchers are scattered among 105 journals. Among them, "Economic Review" secured the first rank for being cited more number of times with 10% of total journal citations, followed by " World Development (5%) and " American Sociological Review (3.28%). The first 09 journals in the rank list contribute nearly 33% of total journal citations. The first 31 journals in the rank list contribute nearly 64% of total journal citations. These 31 journals can be considered as mostly used journals by the sociology researchers.

Conclusion

This study analyzed 12 Masters Theses on Sociology which were submitted to the Main Library, University of Peradeniya during the period of 1995 to 2012. The study discovered that most of the cited sources were books, over 56% of the total citations for graduate works were monographs which were contrasting with previous findings that journal are the most commonly cited format. Since more than 80% of resources cited were 10-30 years old, it meant that postgraduate students prefer to use old materials than current resources. By examining the age of the cited references, data can be used to develop storage and retention policies of the library.

This type of study is also useful in identifying the core journals which are needed for research in sociology. According the findings of present study, it was identified that the journals such as Economic Review, World Development, American Sociological Review, American Journal of Sociology and American Anthropologist were very crucial for postgraduates in sociology in the University of Peradeniya

Consequently, this study should be important in recommending for selection and deselecting of materials. Findings should be relevant to librarians in developing collection development policies and in their budget planning. Examining age of cited references to

help develop storage and retention policies. If the need arises to make cuts to serials budgets and if the collection managers are forced to cancel titles, this data can be used to find the least cited material. This method may also be used if the library needs money to purchase back issues of more heavily used journals.

The results of this study suggest further studies for future research. While journal rankings are usually based on total number of citations, ranking by number of citing authors may be more useful for studies of masters' theses. Further research would be useful to determine if variations in subject resources among disciplines can be generalized with formats other than journals. This study considers only the Sociology masters theses, but this type of studies should be conducted for other subjects too.

References

- Ackerson, L.G. (2001). Is Age an Appropriate Criterion for Moving Journals to Storage. *Collection Management*, 26 (3),63-76.
- Ashman, A. B. (2009). An examination of the research objective of recent citations analysis studies. *Collection Management*, 34(2), 112-128.
- Broadus, R. N. (1977). Citation analysis and library collection building. *Advances in Librarianship*, 7, 300-335.
- Buchanan, A.L. and Herubel, J.P.V.M (1994). Profiling PhD dissertations bibliographies: Serials and collection development in political science. *Behavioral and Social science Librarian*, 13(1), 1-10.
- Budd, J.M. (1986). A citation study of American literature: Implications for collection development, *Collection Management*, 8(2), 49-62.
- Cox, J.E. (2008). Citation analysis of graduate dental theses references: implications for collection development. *Collection Management*, 33(3), 219-226.

- 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.
- Garfield, E. (1977). Is Citation Frequency a Valid Criterion for Selecting Journals? In Essays of an Information Scientist. Philadelphia: ISI Press. Originally published in *Current Comments*, April 5, 1972.
- Georgas, H., & Cullars, J. (2005). A citation study of the characteristics of the linguistics literature. *College and Research Libraries*, 66(6), 496-515.
- Gross, P. L. K., & Gross, E. M. (1927). College libraries and chemical education. *Science*, 66(1713), 385-389.
- Haycock, L. (2004). Citation analysis of education dissertations for collection development. *Library Resources & Technical Services*, 48(2),102-107.
- Johnson, P. (2009). *Fundamentals of Collection Development and Management* (2nd ed). Chicago: American Library Association.
- Kim, M. (2002). Citation Patterns of Korean Physicists and Engineers: Differences by Type of Publication Source and Type of Authorship. *Scienceometrics* 55 (3),421-436.
- Kushkowsky, 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.
- LaBonte, K. (2005). Citation analysis: a method for collection development for a rapidly developing field. *Issues on Science and Technology Librarianship*, 45(2), 23-26
- MacRoberts, M., and MacRoberts, B. (1989). Problems of citation analysis: A critical review. *Journal of the American Society for Information Science*, 40(5), 342.
- Musser, L. R., and Thomas W. C. (1996). Characteristics of Engineering Citations. *Science and Technology Libraries*, 15 (4), 41-49.

- 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. *The Journal of Academic Librarianship*, 33(6), 674-684.
- Rethlefsen, M. (2007). Citation analysis of Minnesota department of health officials publications and journal articles: A need assessment for the RN Barr Library. *Journal of the Medical Library Association*, 95 (3), 260-66.
- Rieb, H.Y. (1993). Citation Analysis: A Case Study of Korean Scientists and Engineers in Electrical and Electronics Engineering. *Proceedings of the ASIS Annual Meeting* ,30,165-171.
- Shafi,S.M. and Gazi, W. (2001) Citation Analysis of PhD Theses : A study of doctoral theses submitted to Kashmir University during 1980-2000 in Natural Sciences, *Trends in Information Management* , 1 (1), 34-39.
- Slutz,M.J (1997). A citation analysis of master theses submitted to the department of English- Kent State University, 457-78.
- Smith, E.T. (2003). Assessing collection usefulness: an investigation of library ownership of the resources graduate students use. *College and Research Libraries*, 64(5),344–355.
- Sylvia, M. (1998). Citation analysis as an unobtrusive method for journal collection evaluation using psychology student research bibliographies, *Collection Building*, 17(1), 20-28.
- Waugh, C. K., & Ruppel, M. (2004). Citation analysis of dissertation, thesis, and research paper references in workforce education and development. *The Journal of Academic Librarianship*, 30(4), 276-284.
- Williams, V. K., & 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*.