The Impact of Subject Heading Assignment on Circulation of Dissertations at Virginia Tech

Richard E. Sapon-White and Mary Hansbrough

Subject headings for bibliographic records for dissertations are no longer assigned at some academic libraries, but the impact this might have on dissertation use has not been evaluated. In this study, bibliographic and circulation records for a sample of 248 academic dissertations were examined to determine the effect of controlled subject headings in the records on circulation of the items. Titles with LC subject headings were compared to those without. Chi-square analysis showed significant differences in circulation for the total sample (p=.04), but not for individual areas of study. Discussion of sources of bias and suggestions for future research are included.

$\mathbf{D}$issertations represent a unique and important form of research literature. Often, they are the first peer-reviewed writings of new researchers and represent the initial appearance of fresh ideas and discoveries in many disciplines. In the United States, over 41,000 dissertations were written during the 1995–1996 academic year, representing degrees awarded at over 750 institutions (UMI 1997). Finding dissertations on a given topic can be challenging because usually they are held only by the library of the institution where the dissertation was written. One way to locate this type of research is through subject searching.

Subject access to theses and dissertations generally takes three forms: published lists, various products published by UMI, and bibliographic records in library catalogs and bibliographic utilities. Lists of notable dissertations in a particular field are sometimes published by professional societies or university departments. These lists are sometimes indexed by standard indexing and abstracting services.

UMI actively microfilms theses and dissertations for subsequent publishing on demand. It publishes its database in a variety of products, such as *Dissertations First, Theses Abstracts,* and *Dissertation Abstracts,* so that researchers can locate and purchase microfilm or print copies of the papers. Subject access to the database is limited as titles are only characterized by subject in the broadest terms. In UMI's electronic products, it is possible to search by keyword.

Librarians in most academic libraries catalog their local dissertations and contribute the bibliographic records to an online bibliographic utility, such as the OCLC Online Computer Library Center, Inc. Online Union Catalog or the Research Libraries Group Research Libraries Information Network (RLIN). The degree to which subject access is provided to this form of literature varies from library to library. Some assign subject headings and classify by subject while others assign no subject headings and place all theses and dissertations in a single class number for the institution.

In recent years, some academic libraries (e.g., University of Florida and UCLA) have chosen to cease assigning subject headings to some theses and dissertations, often as a means of controlling cataloging costs. Another factor in the decision might have been that subject access was thought to be sufficiently provided via keyword searching in the local catalog or by the use of *Dissertation Abstracts*. However, keyword searching is qualitatively different from subject searching using a controlled vocabulary. In addition, the UMI databases can only benefit those searching specifically for theses or dissertations; they do nothing for online catalog searchers seeking subject-specific information owned by that particular library.

The impact on use of dissertations caused by the decision to suspend subject heading assignment remains largely unevaluated. In this paper, we examine the impact on use as evidenced by circulation counts following the implementation of this decision at Virginia Tech.

LITERATURE REVIEW

There is a dearth of literature on the cataloging of theses and dissertations. Much of what does exist is out-of-date and in need of being replicated in an environment of online bibliographic databases and library catalogs. In many of these studies, authors describe how academic libraries catalog theses and dissertations, including how subject analysis is handled (Patterson, White, and Whittaker 1977; Harris and Huffman 1985; Ryans 1991;

Khurshid 1995). Others investigate how researchers locate dissertations (Repp and Glaviano 1987; Lee-Smeltzer and Hackleman 1995) and the ease of searching for dissertations in bibliographic databases (Perry and Salisbury 1995), or evaluate how keyword searching compares with controlled vocabulary searching (Keller 1992). Only in two studies do authors evaluate the impact of subject heading assignment on the circulation of dissertations (Sullivan et al. 1992; Sapon-White 1997).

Patterson, White, and Whittaker (1977) surveyed librarians at 90 university libraries about their local practices for binding, cataloging, classifying, and storing theses and dissertations. They found that 76% of respondents assigned subject headings to all dissertations, while 17% did not assign any subject headings. The remaining respondents assigned only broad subject categories, used locally devised headings, or assigned subject headings on a selective basis. The authors also recorded the statement from one respondent that "subject analysis is too costly for the cataloging department and a waste of time in relation to the amount of use." This unsubstantiated statement was subsequently cited by later authors (Harris and Huffman 1985; Khurshid 1995). It contrasts with the high percentage of libraries reported to be assigning subject headings to dissertations in these and later studies.

Harris and Huffman (1985) reported on a survey of practices at 84 academic libraries concerning thesis and dissertation cataloging practices. They found that 89% assigned subject headings, while 10% did not. Two institutions surveyed, the University of Kansas and the University of New Orleans, had previously performed subject analysis for their dissertations but had discontinued the practice. Harris and Huffman speculated that as libraries switched from card catalogs to online catalogs with keyword searching capabilities, the long descriptive titles of dissertations would obviate the need for subject analysis.

Ryans (1991) surveyed Association of Research Libraries (ARL) libraries in an

attempt to replicate the work of Harris and Huffman (1985). She found that 92% assigned Library of Congress (LC) subject headings, with some respondents also using in-house subject headings.

Khurshid (1995) noted that at his institution, the decision was made to assign only broad LC subject headings followed by a local form subdivision "Theses and dissertations." At his university, the form subject heading "Dissertations, Academic" subdivided by subject was also assigned. He stated that keyword access to the titles and abstracts of theses and dissertations compensated for the lack of more specific subject headings.

Repp and Glaviano (1987) studied how and why extramural researchers (those outside the institution where a dissertation is produced) access locally produced dissertations. They concluded that these scholars usually search for dissertations by subject, though their search terms were often the broad subject categories used by Dissertation Abstracts and related products.

Lee-Smeltzer and Hackleman (1995) surveyed library users, asking those who had checked out a thesis or dissertation about the strategy used to search the online catalog. The authors concluded that library users preferred keyword searching in subject fields over other catalog searching methods, although keyword searching in author fields and exact author searches were also frequently used.

In an interesting study by Perry and Salisbury (1995), OCLC's WorldCat database and RLIN's Dissertation Abstracts International (DAI) database were compared for ease of retrieval of bibliographic records for dissertations. Though they focused primarily on ease of searching and the availability of abstracts, they also compared the databases for coverage of dissertations. They pointed out that there were twice as many records for theses and dissertations in WorldCat as there were in DAI (2.6 million versus 1.3 million) and the WorldCat records appeared much sooner than those in DAI. Moreover, while coverage in the two databases did overlap, coverage was by no means identical. True subject searching in DAI, via a controlled vocabulary, was limited by the broad subject terms used, whereas many of the records in WorldCat were assigned LC subject headings.

Keller (1992) looked at how well the words in subject headings matched words in the titles of theses. Her purpose was to see whether keyword searching could replace subject searching with a controlled vocabulary in an online environment. She found that matching was not high enough to warrant abandoning subject analysis.

Sullivan and others (1992) were one of the few groups of researchers looking at the effect of subject headings on the circulation of dissertations. They looked at circulation records for the first two years of shelf life for dissertations, noting whether or not the titles had circulated. They found that those dissertations that had been assigned subject headings were more likely to have circulated than those that did not have subject headings. In their study, it was impossible to tell how many times an item had circulated.

Similarly, Sapon-White (1997) compared circulation counts for individual titles of 1993 theses at Oregon State University before and after the assignment of subject headings. Circulation counts of titles increased 30% following subject heading assignment. Much of the increase appeared to be due to a small number of titles that circulated many times after the assignment of subject headings.

BACKGROUND

Virginia Tech in Blacksburg is Virginia's land-grant university and its leading research institution. Its research expenditures place it in the top 50 institutions in the nation. Virginia Tech has over 200 degree programs and nearly 25,000 students on campus, including approximately 4,200 graduate students.

The university libraries are a member of ARL and a selective depository for federal documents. The libraries hold over 1.9 million printed volumes, 17,000 serial subscriptions, 5 million microforms, 130,000 audiovisual and machine readable pieces, and 120,000 maps. In addition to the main library, the university

libraries include four branch libraries: Art and Architecture, Geosciences, Veterinary Medicine, and the Northern Virginia Graduate Center.

The university libraries have open stacks and use VTLS as the online catalog system. VTLS is searchable by keyword in almost all text fields, including author, title, series, and all subject heading fields, as well as several note fields (including thesis and summary notes) by performing a "w/" (word) or "b/" (Boolean) search. Keyword searching can also be restricted to the author, title, or subject heading fields if desired. A subject search, "s/," will retrieve only records with LC subject headings in the MARC 650 field.

METHOD

While Sullivan et al. (1992) showed that titles with subject headings were more likely to circulate than those without subject headings, they did not take into consideration the number of times a title circulated. To take an extreme example, it is conceivable that some titles without subject headings circulated many times, while titles with subject headings circulated infrequently. In order to get a complete picture of how subject headings affect circulation, one must consider the number of times a given title is checked out of the library. The present study was designed with this in mind.

The university libraries receive two copies of each dissertation. The first is housed in the Special Collections Department and does not circulate. The second is placed in the general stacks and does circulate. In some cases, a third, circulating copy is deposited in one of the branch libraries. Dissertations are classed together in a single call number and are organized by year, then alphabetically by author's last name. Basic descriptive cataloging is performed by a student assistant using an OCLC constant data form. In the past, a catalog librarian then assigned LC subject headings to each dissertation. After the records are reviewed and edited by a catalog librarian, they are updated in OCLC and exported to VTLS.

Prior to July 26, 1995, LC subject headings were assigned to bibliographic records for all Virginia Tech dissertations. Dissertations cataloged after July 25, 1995, received no LC subject headings, but some were assigned uncontrolled subject terms, suggested by the author of the dissertation, in MARC field 653. Because the VTLS system indexes both title words and field 653 words for general keyword searching, author-supplied terms that matched title words were considered redundant and were omitted from the record during cataloging.

This study was designed to determine the effect of this change in subject heading treatment on dissertation circulation. A program was run against the library database on October 21, 1996, to identify all copies of dissertations that entered into the system between May 1 and October 31, 1995, thus collecting 12 to 18 months worth of circulation data. Entry date was based on the day that the item's barcode was entered into VTLS. This corresponds closely to the date the item becomes available to the public. After being barcoded, the item is labeled and sent to the shelving unit to be placed in the stacks, a process that usually takes only a day or two. The noncirculating Special Collections copies were excluded from the study.

The output of the program listed author, title, call number, barcode number, and copy number for each dissertation that fit the study criteria, as well as the number of times each item had circulated. The circulation count included the number of times the item circulated from the time it entered the library's database until October 21, 1996, the date on which the program was run. Because VTLS updates the circulation count when items are returned, this figure does not include a count for items that were checked out when the program was run or that had been renewed by the borrower after they were checked out. The circulation count does include items checked out to and returned from interlibrary loan.

Following this initial data collection, each bibliographic record was searched in VTLS and viewed to find the date the copy was entered into the system and the academic department for which the dissertation

TABLE 1
ACADEMIC AREAS OF STUDY

Area of Study	Academic Departments
Art/Architecture	Environmental Design and Planning
	Housing, Interior Design, and Resource Management
Business/Management	Business
	Business Administration and Management
	Hospitality and Tourism Management
	Hotel, Restaurant, and Institutional Management
	Management
	Management Science
Education	Administration and Supervision of Special Education
	Adult and Continuing Education
	Community College Education
	Counseling/School Psychology
	Counseling and Student Personnel Services
	Counselor Education
	Curriculum and Instruction
	Educational Administration
	Educational Research and Evaluation
	Instructional Technology and Design
	Vocational and Technical Education
Engineering	Aerospace and Ocean Engineering
	Aerospace Engineering
	Agricultural Engineering
	Biological Systems Engineering
	Chemical Engineering
	Civil Engineering
	Electrical Engineering
	Engineering Mechanics
	Environmental Engineering
	Industrial and Systems Engineering
	Materials Engineering Science
	Mechanical Engineering
	Mining and Minerals Engineering
Science	Animal and Poultry Science
	Animal Science
	Biochemistry and Anaerobic Microbiology
	Biology
	Botany
	Chemistry
	Computer Science
	Crop and Soil Environmental Science
	Fisheries and Wildlife
	Food Science and Technology
	Forestry
	Genetics

TABLE 1 (cont.)
ACADEMIC AREAS OF STUDY

	Horticulture
	Human Nutrition and Foods
	Mathematical Physics
	Mathematics
	Physics
	Statistics
	Veterinary Medical Science
	Wood Science and Forest Products
Social Science	Agricultural and Applied Economics
	Clinical Psychology
	Family and Child Development
	Psychology
	Public Administration/Public Affairs
	Science and Technology Studies

was written. If uncontrolled subject terms were encountered in a 653 MARC field, they were counted and recorded for later analysis.

The circulation count for dissertations with subject headings was then compared to the circulation count of those cataloged without subject headings. Data were entered into an Excel spreadsheet for analysis. Each entry was assigned to one of six broad areas of study (art and architecture, business and management, education, engineering, science, social science). The information was sorted by academic area so that the data could be analyzed by subject categories as well as for the whole sample. The areas of study and the departments they included are listed in table 1.

Anomalies in the data included 6 titles with 2 copies that could be checked out. In 5 of these 6 cases, 1 of the 2 copies did not circulate at all. In all 6 cases, the circulation counts for the 2 copies were added together and entered as if there were only 1 copy. Also, 2 dissertations were encountered that comprised 2 volumes each. In one case, neither volume circulated. In the second case, 1 volume circulated once while the other did not circulate at all. Again, the circulation information was added together and entered as if it were one volume.

Although titles with subject headings were on the shelf for a longer period of time than those without subject headings (potentially 15–18 months versus 12–15 months), and therefore had a greater likelihood of circulating, this is probably not a significant factor for two reasons. First, the extra months occurred in the summer, which is a time when the library receives the lowest amount of foot traffic and circulation of dissertations in both groups would be expected to be extremely low. Second, although the date parameters for selection of the sample cover a 6-month period (3 months before and after the change in treatment), in actual fact 92% of the titles with subject headings entered the online system between June 21 and July 25, 1995; 58% of titles without subject headings entered the system between July 26 and August 15, 1995. Any bias as a result of the difference in time on the shelf was therefore considered to be very small.

Data were analyzed using the chi-square test of independence (Sheskin 1997). Because Sullivan et al. (1992) had shown that titles with subject headings were more likely to circulate, this study was designed to evaluate whether titles with subject headings circulated more often than those without subject headings. The null hypothesis was: The circulation count for dissertations is independent of the subject headings assigned to their

TABLE 2
DISTRIBUTION BY ACADEMIC AREA AND SUBJECT HEADING TREATMENT

Academic Area	With Subject Headings	Without Subject Headings	Total
Art/Architecture	1	2	3
Business/Management	5	4	9
Education	40	24	64
Engineering	51	28	79
Science	31	34	65
Social Science	11	17	28
Total	139	109	248

bibliographic records. A value of .05 was chosen for determining significance of the results.

RESULTS

A total of 248 dissertations were cataloged between May 1, 1995, and October 31, 1995. Of these, 139 were assigned LC subject headings. Of the 109 without subject headings, 36 had author-supplied uncontrolled subject terms added in the MARC 653 field of the record. Although the program parameters were set to collect 6 months worth of cataloging data, in actuality, all of the dissertations in the study were cataloged between June 21, 1995, and October 31, 1995. Therefore, the data collected represents 16–17 months worth of circulation information for dissertations with LC subject headings and 12–15 months worth of information for those without subject headings.

The distribution of the records among areas of study is shown in table 2.

The mean number of circulations for the total sample was 58% higher for records with subject headings when compared to those without subject headings. Engineering records showed a mean 69% higher for records with subject headings, while Science records with subject headings had a mean 128% higher than those without subject headings. Education records, however, showed nearly identical means for the two groups.

The number of titles in Art/Architecture, Business/Management, and Social Science was too small to analyze using the chi-square test but are included in the overall analysis. Results of the test are shown in tables 3–6 for all dissertations, Education, Engineering, and Science, respectively. Statistically significant results were found for the aggregated sample (chi-square=9.99; df=4; p=.04), but not for Education (chi-square=1.27; df=2; p=.53), Engineering (chi-square=3.74; df=2; p=.15), or Science dissertations (chi-square=1.14; df=1; p=.28).

DISCUSSION

The results shown in table 3 indicate that, for all titles in this study, dissertations with subject headings were more likely to circulate and circulated a greater number of times than titles without subject headings. The null hypothesis that circulation count is independent of the presence of subject headings is therefore rejected. One explanation for these results is that library patrons used subject searching in the online catalog and found their way to dissertations, as well as other materials, on their subject of interest. Conventional subject searching retrieved materials that had subject headings on their records, while records lacking subject headings were not retrieved. In the case of keyword searching, the addition of the subject headings might have enriched the vocabulary present in the bibliographic record, thereby increasing the likelihood of subject retrieval by keyword. These results indicate that library users are searching by subject often enough to have an impact on the circulation of dissertations.

One might reasonably expect that subject terms supplied by the dissertation

TABLE 3
ALL DISSERTATIONS

	Number of Circulations						
	0	1	2	3	>3	Total	Mean
Observed							
Without subject headings	58	29	15	7	0	109	0.73
With subject headings	53	42	24	13	7	139	1.15
Total	111	71	39	20	7	248	
Expected							
Without subject headings	48.79	31.21	17.14	8.79	3.08	109	
With subject headings	62.21	39.79	21.86	11.21	3.92	139	
Total	111	71	39	20	7	248	

Chi-square=9.99; p=0.04

authors (and recorded in the 653 MARC field) would boost the circulation of titles without subject headings. In the present study, any such boost was not great enough to be detected. This might be due to the fact that only 36 titles of the 109 without subject headings (33%) had uncontrolled subject terms added to the bibliographic record. In order to fully evaluate the effect on circulation of adding uncontrolled subject terms, one would need to compare a group of records containing such terms with a control group lacking both author-supplied terms and LC subject headings.

These results help to confirm the work done by Sullivan et al. (1992). In that study, researchers could only tell whether a dissertation circulated or not; there was no way to account for the number of times a title circulated. In the present investiga-

tion, by looking at the circulation counts for each title, we show quantitatively higher circulation counts for dissertations with subject headings.

The results from the academic fields of study tables are less revealing as there were no statistically significant differences between titles with and those without subject headings in any disciplinary area. The chi-square test, however, is not powerful enough to detect differences in the distribution of circulation counts when the sample size is small. A more sophisticated test, beyond the means available for the present study, might have the power to detect a directional shift in these tables. That shift is evident if one compares the observed and expected values within the engineering or science tables; titles with subject headings circulated more than expected while titles without

TABLE 4
EDUCATION DISSERTATIONS

	Number of Circulations				
	0	1	>1	Total	Mean
Observed					
Without subject headings	6	6	11	23	1.26
With subject headings	9	16	15	40	1.25
Total	15	22	26	63	
Expected					
Without subject headings	5.48	8.03	9.49	23	
With subject headings	9.52	13.97	16.51	40	
Total	15	22	26	63	

Chi-square=1.27; p=0.53

TABLE 5
ENGINEERING DISSERTATIONS

	Number of Circulations				
	0	1	>1	Total	Mean
Observed					
Without subject headings	15	5	8	28	0.89
With subject headings	16	13	22	51	1.51
Total	31	18	30	79	
Expected					
Without subject headings	10.99	6.38	10.63	28	
With subject headings	20.01	11.62	19.37	51	
Total	31	18	30	79	

Chi-square=3.74; p=0.15

subject headings circulated less than expected. The need to "collapse" the cells to perform the chi-square analysis, partially due to the small sample size, is the likeliest reason why no statistically significant probability is calculated for the science and engineering tables. The aforementioned directional shift is clearly missing from the education table.

Harris and Huffman (1985) contended that keyword searching might provide a useful substitute for subject searching when titles are sufficiently descriptive of content. In this study, all titles were highly descriptive of their content. Though some dissertations in the humanities have non-descriptive, even whimsical, titles, none such works were present in this study. In the case of whimsical titles, retrieval by keyword in the absence of subject heading assignment would be highly unlikely.

These results have implications for understanding user behavior. Clearly, dissertation users are not only conducting known-author or known-title searches. Subject access to dissertations is important, though one cannot tell from this study whether the catalog searchers are specifically looking for dissertations or if they are searching for material on their subject in any format. As previously mentioned, authors of two studies have looked at this issue and, in both instances, concluded that, for dissertation users, searching by subject is significant (Repp and Glaviano 1987; Lee-Smeltzer and Hackleman 1995).

We have used circulation counts as an indication of use, but do not consider in-house use. A more complete picture of the effect of subject heading assignment on use would have to include such use as

TABLE 6
SCIENCE DISSERTATIONS

	Number of Circulations			
	0	>0	Total	Mean
Observed				
Without subject headings	27	7	34	0.21
With subject headings	21	10	31	0.48
Total	48	17	65	
Expected				
Without subject headings	25.11	8.89	34	
With subject headings	22.89	8.11	31	
Total	48	17	65	

Chi-square=1.14; p=0.28

well. It does not seem likely that in-house users would behave differently from users who check out materials, and so the results here are not likely to be biased by the focus on circulation counts.

Though beyond the scope of the present study, the question still remains as to whether subject searching in *Dissertation Abstracts* could compensate for the lack of subject headings assigned to dissertations in a local catalog. It would be interesting in a future study to compare the descriptors used in DAI's subject searching with the controlled vocabulary of LC subject headings in a library's catalog. Such a study would shed further light on the merits of these two subject searching systems and add to our understanding of how users find the materials they need.

Further research, perhaps using a larger sample and a more powerful statistical test, is indicated to evaluate differences between academic areas of study and to corroborate these results with an assessment of user behavior when searching for dissertations.

CONCLUSION

Cost-saving measures in academic libraries must be carefully evaluated in terms of their effect on service to library users. One such measure, the suspension of assigning subject headings to dissertations, has a significant impact on dissertation use. The results of this study indicate that dissertations with subject headings in their bibliographic records are more likely to circulate (and circulate more often) than those without subject headings. Failure to add subject headings to bibliographic records for dissertations, while indeed saving time in cataloging, noticeably reduces the likelihood that the items will be found and used by patrons. Library administrators choosing to make such a decision should do so understanding that there will be a resulting loss of use of this unique form of research literature.

WORKS CITED

Harris, George, and Robert Huffman. 1985. Cataloging of theses: A survey. *Cataloging & classification quarterly* 5, no. 4: 1–15.

Keller, Barbara. 1992. Subject content through title: A masters theses matching study at Indiana State University. *Cataloging & classification quarterly* 15, no. 3: 69–80.

Khurshid, Zahiruddin. 1995. Improvisations in cataloging of theses and dissertations. *Cataloging & classification quarterly* 20, no. 2: 51–59.

Lee-Smeltzer, Janet, and Deb Hackleman. 1995. Access to OSU theses and dissertations in Kerr Library: How they are used . . . or are they? *Cataloging & classification quarterly* 12, no. 4: 25–43.

Patterson, Kelly, Carol White, and Martha Whittaker. 1977. Thesis handling in university libraries. *Library resources & technical services* 21: 274–285.

Perry, Stephen, and Lutishoor Salisbury. 1995. Access to information in both CitaDel and FirstSearch: A comparative study of dissertation coverage. *Information technology and libraries* 14, no. 1: 17–29.

Repp, Joan M., and Cliff Glaviano. 1987. Dissertations: A study of the scholar's approach. *College and research libraries* 48: 148–59.

Ryans, Cynthia C. 1991. Cataloging theses and dissertations: An update. *Cataloging & classification quarterly* 14, no. 1: 83–87.

Sapon-White, Richard. 1997. Impact of subject heading assignment on circulation of theses at Oregon State University. Poster session presented at the annual meeting of the American Library Association, San Francisco, Calif.

Sheskin, David J. 1997. *Handbook of parametric and nonparametric statistical procedures.* Boca Raton, Fla.: CRC Press.

Sullivan, Michael V., and others. 1992. Effects on circulation of dissertations of omitting subject headings. In *Academic libraries: Achieving excellence in higher education,* ed. Thomas Kirk. 122–27. Chicago: Association of College and Research Libraries.

UMI. 1997. *American doctoral dissertations, 1995–1996.* Ann Arbor, Mich.: UMI.