July 2007

A Rough Measure of Copy Cataloging Productivity in the Academic Library

John Buschman
Rider University, buschman@rider.edu

F. William Chickering
Rider University, wchickering@rider.edu

Library Philosophy and Practice 2007

ISSN 1522-0222

A Rough Measure of Copy Cataloging Productivity in the Academic Library

John Buschman
Professor-Librarian and Chair

F. William Chickering

Dean of University Libraries
Rider University Library
2083 Lawrenceville Road
Lawrenceville, New Jersey 08648

Introduction

Cataloging productivity has been notoriously difficult to measure. This is borne out by the fact that, while enough has been written to produce two bibliographies on the subject (in 1970 and 1988 - and more since), the vast majority of this work "is fragmentary, limited in scope, and short on detail (Morris, Herbert, Osmus, & Wool, 2000). The research is near-impossible to replicate because it is based on locally produced (and undefined) data, it is focused often on the productivity of individuals in a local circumstance, and it is frequently cast in terms of costs and benefits where costs are often difficult to define and measure and thoroughly monetizing the benefits of libraries has been largely absent (Charbonneau, 2005; University of Maryland Libraries, 2001). The issue has frequently boiled down to the "allegorical battle between Quantity and Quality" (Mandel, 1988, p. 215) with much room for posturing over needs for justification from management and the indignant reply of placing monetary value on such a thing. Hence productivity data is "scanty" and tends to focus on turn around time - a measure itself going back thirty years in ARL libraries (University of Maryland Libraries, 2001).

In turn, much has been said over a long period of time about the death of the book: McLuhan's (1962; 1965; see also Neill, 1971) work implied that t he book was dead – superseded by media culture; and F.W. Lancaster (1982, p. 149-150) predicted about a quarter century ago that, as the “print on paper literature of the past” is replaced by electronic versions, we will see “the disappearance of the library… All that will remain are a few institutions that preserve the printed records of the past” – hence “the paperless society.” Despite this, in 2004 (a year of relevance in this article), 375,000 new titles and editions of books were published in the English speaking world, and in

terms of U.S. book production, the numbers of titles have grown substantially over the last decade from 104,124 in 1993 to 190,078 in 2004 (Bowker.com, 2005; Grabois, 2005). Academic libraries (both the largest and smaller ones) have in turn increased expenditures and acquisitions of books (Stoller, 2006), so it is still a reasonable assumption for the foreseeable future that academic libraries will acquire a not insignificant portion of this output.

Rider University Library has followed this general pattern. A private, coeducational university located in New Jersey, Rider has two campuses and libraries (Lawrenceville and Princeton), 3,764 full-time undergraduates, 822 part-time, and 1,204 graduate students. The Moore Library in Lawrenceville is a medium sized academic library: 450,000 volumes, 80 databases, and 20,000+ journal titles - mostly in electronic formats, 10 library faculty and 13 support staff with 2 supervisors. In 2003, a major review of spending patterns and curriculum support needs resulted in a significant reconceptualization and shift in collection philosophy at Moore Library. Discontinuing some print duplication of stable electronic periodical content resulted in the ability to shift needed funding into monographs acquisition. This strategy was augmented by a long overdue increase in the monographs collection budget. The result was an increase in monographs acquisition, from 3311 purchased volumes in AY 2000-2001, to 5029 purchased volumes in AY 2004-2005 - an almost 52% increase. Total added monographic volumes went from 5173 to 6474, representing an approximately 25% rise. It is due to this increase and a related problem that we produced a rough copy cataloging productivity benchmark. It is worth noting here that by benchmark we mean primarily quantity of work, rather than the concept of "'production benchmarks'... developed to measure and compare quantitative and qualitative output" primarily applied to professional catalogers working at higher skill-set levels (Charbonneau, 2005, p. 41). Faced with a management challenge in terms of both an increase in monographic acquisitions and recurring requests for additional copy cataloging help, questions concerning the actual need for that help arose.

The Need for a Productivity Benchmark

A literature search was performed, but no national standards or benchmarks were discovered in a series of standard library school texts - current and retrospective - on cataloging relevant to an academic context (Taylor, 1988; System and Procedures Exchange Center, 1987; Taylor, 2000; Rowley & Farrow, 2000; Intner & Hill, 1989; Intner & Fang, 1991). Those that did give a nod towards the subject made broad statements on outsourcing not necessarily saving money (Wilson & Colver, 1997), descriptions of what performance standards are and what they measure, why they are important, and the resistance of staff to these efforts (Foster, 1987, p. 124-129; University of Maryland Libraries, 2001; ALCTS Technical Services Cost Committee, 1991), that copy cataloging was now a job for paraprofessionals (Rhee, 1998; Evans & Heft, 1994, p. 209), or focus on unit costs for acquisitions (Bierman, 1989; Rebarcak & Morris, 1996). Kohl (1986, p. 22-23) pulled together relevant studies over twenty years ago and found interesting data on hit rates and the quality of cataloging copy available, but no productivity benchmarks. Charbonneau (2005, p. 44) produced a recent overview

of the situation, and confirmed this observing that “no blanket benchmark exists that can refer to all catalogers or groups of catalogers. The process becomes one of establishing individual benchmarks based on each cataloger's responsibilities and situation.” McCain and Shorten (2002, p. 23, 27) presented findings of survey results of 27 Association of Research Libraries (ARL) members. The heart of the matter from a management (and budgeting) point of view quickly became clear: “Since data on costs are not necessarily comparable among institutions, other quantifiable measures of efficiency and effectiveness would enhance managerial decision-making.” An illustration of the problem appears in their comparisons of catalog updates per full time equivalent employee (FTE) for FY 1998-99. The reported figures are “an average of 4,057, a median of 635, and a standard deviation of 8,132.” Such statistical disparities provide little help in moving toward a practical benchmark. In large part the problem of comparability lies in the myriad ways libraries combine elements from a large menu of possible technical services operations performed with a variety of degrees of tolerated quality variations. One result of this review was the insight that cost studies contained a wealth of data on cataloging and time spent on cataloging. This would prove valuable in the course of this inquiry.

A Rider Benchmark?

At Rider's main library, copy cataloging is a soup-to-nuts process done by non-professional support staff:

- reviewing OCLC records used to order the material and/or searching for appropriate OCLC records to match the added volume within defined parameters;
- importing them into the local database and representing those holdings in OCLC;
- minor modifications thereto according to written guidelines produced by the primary catalog librarian and/or referral to a cataloger assigned to review, note corrections needed-on-the-fly and return to the staff for completion;
- deleting records and modifying OCLC holdings as appropriate;
- systematic minor catalog corrections as instructed (for records of classes of materials like DVDs, for instance, or information from inventory reviews);
- physical processing of the books with supplemental help from student workers for this last step under standards and guidelines set up by the library administration.

Long-term decisions on acceptable standards and sources of records, use (or non-use) of subfields, etc. and the resulting access points and policies are in the hands of library faculty and administration. Quality of the output is overseen by periodic review from the staff supervisors, department chair, and the library faculty. Copy cataloging at Rider is therefore a thoroughly documented *process* to be followed, with the exceptions referred to a librarian to resolve (quickly). “How much does copy cataloging cost per volume?” was not the right question since information was produced on the costs for importing records and on processing materials, but not for staff time in any useful per unit version. What really was lacking was benchmark information on the *rate* at which copy catalogers can *be anticipated* to do copy cataloging. With this missing piece of

data, it would be easily possible to take part-time and full time pay rates and create cost models for dealing with the increased need for copy cataloging support and/or measure the productivity of existing staff - which became an increasing focus as our investigations proceeded.

In our concern over (perhaps) perceived differences in productivity, we reviewed the raw data available from the cataloging subsystem, and what we came up with was the concept of the number of records “touched” (that is, modified, added, deleted, etc.) in aggregate. This has the advantage of accounting for *all* of the variety of work our copy cataloging staff do with records, both simple and complex. Second, we were careful to review this aggregated statistic over time. That is, no day-to-day or week-to-week measure could possibly be fair, since any *particular* period of work time may be impeded by the onset of a cold, bad news from home, a set of difficult records for the books, etc., or conversely, inflated by a very easy set of records or the return of a librarian who can quickly resolve dozens of questions - resulting in a seeming spike in productivity. Viewing this data over time also insulated us from the charge of micromanaging or unfair micromeasuring and gave staff copy catalogers the benefit of looking at their overall production, not just a few snapshots. Finally, we produced a calculation of time actually at work during the periods corresponding to the online system reports: the total, possible days which could be worked during that period (e.g. not counting holidays or snow days when the university was closed), subtracted days not worked by each copy cataloger (vacation, illness, etc.), times the number of hours per day to be worked for an individual calculation for each copy cataloger as to the actual hours worked over a significant period.

With this information in combination, it was possible to compare a rough measure of productivity of copy cataloging staff in gross terms. The results of copy cataloging activity (records "touched") were divided by the hours worked to determine a rough measure of productivity - the number of records "touched" per hour worked. The resulting measure emerged with startling individual differences:

- From 7-1-02 to 6-30-03 staff member A "touched" 2133 catalog records;
- Staff member B "touched" 1736 over the same period.
- Staff member A worked a total of 1544.5 hours from 7-1-02 to 6-30-03;
- Staff member B worked 1380 hours over the same period.

Given the amount of hours worked vs. the records "touched" in the system staff member A "touched" 1.38 catalog records/hour worked and staff member B "touched" 1.25. This is a 10.5% differential in productivity - which got worse:

- From 7-1-03 to 6-30-04, staff member A "touched" 2363 catalog records;
- Staff member B touched 1710 over the same period.

During this time period, staff member A worked a total of 1633.5 hours and B worked 1448 hours. This results in a differential of 1.45 records/hour worked (for A) vs. 1.18 (for B) - now a 22.8% differential - and a drop between the two periods - in

productivity. Anecdotally, we knew the issue was time on task, and a working assumption on the part of staff member B that staff member A would do the work she/he did not get done. It is worth noting here that the basic job descriptions for both are identical. However, staff member A compiles statistics, is a work leader, directs student workers, and handles the more difficult records (microform, A-V, Reference, and "Rush" items). That is, those records A "touches" are far more complex and there are ancillary issues as part of staff member A's work. Furthermore, we made an adjustment midway through to boost staff member B's productivity by removing all physical processing responsibilities and limited the scope of the types of materials in order to enable consistent concentration only on one type of copy cataloging. Our preliminary conclusions were clear: it was reasonable to expect some consistent level of copy cataloging productivity per hour, and this number, averaged over time should be relatively consistent for the individual, and within a modest band, between individuals. However, standards that go beyond comparisons of individual differences are needed when personnel issues are at stake. The bottom line question was always: is the expectation to approach 1.4 records "touched" per hour fair? We went back again to some of the cost studies to see if similar data could be extracted.

A Re-review of Data

Some larger institutions had undertaken cataloging cost studies, but we approached that data informed by our own calculations, and three were the most thorough and useful.

1. Vanderbilt University Library's Staff Time Allocation Study (2003) and their ARL Statistics (2003/2004; 2004/2005) lists copy cataloging tasks similar to Rider, the cumulative cataloging staff work week for the study (431 hours), and monographs purchases for similar time periods (26,592 for 2004/05 and 33,497 for 2003/04). The monographs figures are a stand-in for our records "touched" definition given that some would have been diverted for original cataloging, there are separate physical processing and catalog maintenance centers at Vanderbilt, but items withdrawn in any given year should account for the absence of time spent on those tasks. As a rule-of-thumb, attendance rate (total hours actually worked out of the total hours available to work) is assumed to be 85%: national data calculate paid time off as a cost of 10-11% of payroll, and the education field has a 4% absentee rate (Cornell University Libraries, Catherwood Library, 2005). Therefore Vanderbilt University Library's copy cataloging work year would comprise about 19,050 hours, and the copy cataloging staff's overall productivity is an average rate of 1.59 records/hour over both years - similar to our own upper and, we felt, reasonable measure of productivity.
2. McCain and Shorten (2002) surveyed 27 ARL libraries to measure "efficiency and effectiveness." Many variables were not comparable with Rider, but Table 5 (p. 29) on cataloging efficiency gave a ranked list of volumes cataloged per FTE. Taking the median ranking of the 25 institutions who reported the figure, 2,674 volumes cataloged per FTE could then be calculated against an average academic library work year in total hours. (For reference, the average among

these 25 institutions is 2,681 volumes cataloged/year.) That is, 52 weeks times 5 days per week = 260 possible work days, multiplied times 8 hours per day and the 85% rule-of-thumb attendance. This comes to 1,768 hours worked per year by an FTE - or 1.51 volumes per hour cataloged. Again, this comes close to the Rider and Vanderbilt rough productivity standard. The justification for using the volumes figure (vs. titles) is that the activity on the added volumes would more closely correspond to a Rider record "touched." Again, it is in the aggregate measured over some time that the rough measures fall within some reasonable ranges.

3. Lastly, Iowa State University Library updated an earlier study of cataloging costs (Morris, 1992) in 2000 (Morris et al). Like the other two studies, the tables were filled with data presenting the same problem and opportunity on transferability of the results. For instance, the summary section on copy cataloging per-title costs notes an average of $12.22 for a monograph (e.g. no other formats and no substantive professional's necessary involvement needed) for 1997/98. This data was less than useful since it is not entirely clear if this particular figure includes benefits, overhead, etc. and the staff salary information provided for the period ranged from $21,600 to $34,700 per year. This means that copy cataloged books based on this salary range (without benefits, etc. calculated in) would be done at a rate of 1.18/hour by the lowest paid staff (who would make about $10.40/hour), and a rate of.73/hour by the highest paid staff. This was not comparable data. However, the study reported 17,809 DLC volumes copy cataloged requiring no cataloger review (Table 4) in 196 average hours per week of copy cataloging (Table 2) for 1997/98. This comes out to 1.75 records/hour. Like Vanderbilt, Iowa State University Library has separate centers for volume preparation, preservation, and catalog maintenance. We would argue again that this productivity rate falls within a reasonable range of the levels of expected productivity we calculated at Rider.

Conclusion: The 3000 Foot Level Analysis

While our review of our own and other academic libraries' data does not provide conclusive statistical "proof" of a copy cataloging productivity benchmark, we would argue that it is a highly suggestive means of approaching the problem and a productive avenue for further study. It is reasonable to posit that the sheer amount of details within technical services have sometimes gotten in the way when this issue has been studied, resulting in the noted problems between and among the studies. Data reviewed from 3000-foot level is perhaps more comparable and revealing that those with too much specific detail. When one steps back from so many documented particulars and looks at aggregates of materials copy cataloged/records "touched" over longer periods of time measured against hours worked, a rough benchmark of copy cataloging output emerges of about 1.4 to 1.7 volumes/records per hour. This is a fair measure allowing for good weeks and bad weeks, easy materials and hard, etc. - all measured over a substantial amount of time worked to ensure fairness.

On a local level, analysis of all available data led to the conclusion that staff member B was grossly under producing - a 28% negative differential that further data indicated was growing in the intervening months, and a change was made to bring in another staff member. Two post-tests were conducted. First, Staff member A was asked to finish the work left over after the departure of staff member B (to see if the work assigned was similar in difficulty, and therefore could reasonably be done at the same rate), and productivity was measured. It again fell within the 1.4 to 1.7 records/hour "touched" range. After the new staff member was hired and trained, data again indicated that this new person was performing at about the productivity level of staff member A. Lastly, not only did this process allow Rider University Library to create measures for copy cataloging productivity, it also allowed a process of validation of real staff achievement.

References

ALCTS Technical Services Cost Committee. (1991). Guide to Cost Analysis of Acquisitions and Cataloging in Libraries. *ALCTS Newsletter*, *2* (5), 49-52.

Bierman, K. J. (1989). What Does It Cost to Acquire a Book? The Tucson Public Library Expereience. In G. M. Pitkin (Ed.), *Cost Effective Technical Services: How to Track, Manage and Justify Internal Operations.* New York : Neal-Schuman.

Bowker.com. (October 12, 2005). *English-Speaking Countries Published 375,000 New Books Worldwide in 2004.* Retrieved March 27, 2006, from http://www.bowker.com/press/bowker/2005_1012_bowker.htm.

Charbonneau, M. D. (2005). Production Benchmarks for Catalogers in Academic Libraries: Are We There Yet? *Library Resources & Technical Services, 49* (1), 40-48.

Cornell University Libraries, Catherwood Library. (July 2005). *Reference Question of the Month: What is the Average Rate of Employee Absenteeism?* Retrieved March 1, 2006, from http://www.ilr.cornell.edu/library/research/questionofthemonth/jul05.html.

Evans, G. E., & Heft, S. M. (1994). *Introduction to Technical Services* (6th ed.). Englewood, CO : Libraries Unlimited.

Foster, D. L. (1987). *Managing the Catalog* Department. Metuchen, N.J. : Scarecrow Press.

Grabois, A. (Comp.). (2006) *U.S. Book Production, All Hardback and Paperback, 1993-2005 Preliminary.* Retrieved March 27, 2006, from http://www.bookwire.com/bookwire/decadebookproduction.html.

Intner, S. S. & Hill, J. S. (Eds.). (1989). *Recruiting, Educating, and Training Cataloging Librarians: Solving the Problems.* New York : Greenwood Press.

Intner, S. S. & Fang, J. R. (1991). *Technical Services in the Medium-sized Library: An Investigation of Current Practices.* Hamden, CT : Library Professional Publications.

Kohl, D. F. (1986). *Cataloging and Catalogs: A Handbook for Library Management.* Santa Barbara, CA : ABC-Clio.

Lancaster, F. W. (1982). *Libraries and Librarians in an Age of Electronics.* Arlington, VA: Information Resources Press.

Mandel, C. A. (1988). Trade-offs: Quantifying Quality in Library Technical Services. *Journal of Academic Librarianship, 14* (4), 214-220.

McCain, C. & Shorten, J. (2002). Cataloging Efficiency and Effectiveness. *Library Resources & Technical Services 46* (1), 23-31.

McLuhan, M. (1962). *The Gutenberg Galaxy: The Making of Typographic Man.* Toronto : University of Toronto Press.

McLuhan, M. (1965). *Understanding Media: The Extensions of Man.* New York : McGraw-Hill.

Morris, D. E. (1992). Staff Time and Costs for Cataloging at Iowa State University. *Library Resources & Technical Services 36,* 79-95.

Morris, D. E., Herbert, C. B., Osmus, L., & Wool, G. (2000). Cataloging Staff Costs Revisited. *Library Resources & Technical Services, 44* (2), 70-83.

Neill, S. (1971). Books and Marshall McLuhan. *Library Quarterly 41* (4), 311-319.

Rebarcak, P. Z., & Morris, D. (1996). The Economics of Monographs Acquisitions: A Time/Cost Study Conducted at Iowa State University. *Library Acquisitions: Practice & Theory 20* (1), 65-76.

Rhee, S. F. (1988). Technical Services Budgeting and Finance. In M. Gorman (Ed.), *Technical Services Today and Tomorrow* (2nd ed.). Englewood, CO : Libraries Unlimited.

Rowley, J. & Farrow, J. (2000). *Organizing Knowledge: An Introduction to Managing Access to Information.* Burlington, VT : Gower.

Stoller, M. (2006). A Decade of ARL Collection Development: A Look at the Data. *Collection Building 25* (2), 45-51.

System and Procedures Exchange Center. (1987). *SPEC Kit 136, Managing Copy Cataloging in ARL Libraries.* Washington, D.C. : Association of Research Libraries, Office of Management Studies.

Taylor, A. G. (1988). *Cataloging with Copy: A Decision-maker's Handbook.* Englewood, CO : Libraries Unlimited.

Taylor, A. G. (2000). *Wynar's Introduction to Cataloging and Classification.* Englewood, CO : Libraries Unlimited.

University of Maryland Libraries. (March 15, 2001). *Technical Services Program Review, Benchmarking Group Report.* Retreived March 27, 2006, from http://www.lib.umd.edu/TSD/PROGRAMREV/benchmarking.html.

Vanderbilt University Library. (November 10-16, 2003). *Staff Time Allocation Study.* Retrieved January 16, 2007, from http://staffweb.library.vanderbilt.edu/admin/TimeStudy/VanderbiltTestData_files/frame.htm.

Vanderbilt University Library. (2003/2004 - 2004/2005). *ARL Statistics*. Retrieved January 16, 2007, from http://staffweb.library.vanderbilt.edu/admin/arl0405.htm.

Wilson, K. A. & Colver, M. (Eds.). (1997). *Outsourcing Library Technical Services Operations: Practices in Academic, Public, and Special Libraries.* Chicago: American Library Association.