

Research Notes

CONTENTS

<i>Pal V. Rao</i>	341	The Relationship between Card Catalog Access Points and the Recorded Use of Education Books in a University Library
<i>Mel Westerman</i>	346	Salary Comparisons between Academic Librarians and Instructional Faculty

PAL V. RAO

The Relationship between Card Catalog Access Points and the Recorded Use of Education Books in a University Library

HYPOTHESES

This research study is concerned with the investigation of any statistically discernible relationship between the number of card catalog access points provided for a group of randomly selected book titles in an academic library and the number of times the same titles circulated in a specified period of time.

The following hypotheses are advanced:

H_0 : There is no significant relationship between the number of card catalog access points per title and recorded use per title in an academic library for a specified period of time.

H_1 : There is a significant relationship between the number of card catalog access points per title and recorded use per title in an academic library for a specified period of time.

PURPOSE

There appear to be very few studies con-

cerned with the relationship between card catalog access points and other components of library operations. McClure stated, "The paucity of research becomes quite apparent after one checks *Library Literature* and quickly discovers that the subject headings 'added entries,' 'subject added entries,' or 'tracings' are not used. Information about the number and types of subject and added entries have appeared largely as an aside to other topics being researched."¹

The purpose of this study is to make a small contribution in this area and to determine whether any relationship exists between card catalog access points and library circulation.

LITERATURE REVIEW

As indicated by McClure,² there has been very limited research in this area. McClure examined the trends of assigning added entries from 1950 to 1973. He found that the average number of subject-added entries per card was 1.2 in 1950, and it increased to 1.3 in 1973. Similarly, the average number of added entries per card increased from 1.0 in 1950 to 1.5 in 1973. The results of this study

Pal V. Rao is professor and head of Information Systems Department, Booth Library, Eastern Illinois University, Charleston.

indicate that there was a definite trend for more added entries in 1973 than in 1950. Obviously, more cards mean greater costs for filing time and card production and storage. What are the benefits of these additional costs? McClure's study does not address that question but suggests the need for further research to identify the effects of these additional cards on the library operations.

In addition to the McClure study, there are a number of literature sources that are applicable on a partial basis to the present study and will be referenced here at appropriate points.

LIMITATIONS

Although this study is concerned with a basic correlation between two variables, the correlation itself does not indicate any causality. Consequently, the findings of this study should not be interpreted as causal links.

There is a need for research that establishes interrelationships among various components of a library system. It is hoped that, at some future time, we will have all the relevant interrelationships established. This study attempts to make a small start towards that eventual goal.

DEFINITIONS

Card catalog access point: an entry in the Eastern Illinois University Library dictionary providing bibliographic access to a publication through author, title, series, subject headings, and analytical approaches.

Recorded use: an electronic record of materials checked out from the Eastern Illinois University Library's circulating collection.

RESEARCH METHOD

Data Collection for the Selected Variables

Number of Card Catalog Access Points. For each selected sample title, the number of added entries provided on the main entry was recorded. Whether all the added entries on the main card actually appear in the card catalog was also examined. Thus a number of catalog access points (i.e., main entry plus added entries) for each selected title were obtained.

Number of Circulations for a Given Period. Booth Library has been operating a

computerized circulation system since September 1968. A by-product of this system is a complete circulation transaction file, which contains a separate record for each circulation transaction conducted by the library. Transactions conducted while the computer was down were subsequently updated using a batch update procedure. Consequently, the library has a complete circulation transaction file since September 1968 on archival tapes.

Each record on these tapes contains the following data elements: (1) book identification; (2) date of transaction; (3) time of transaction; (4) patron identification (social security number); (5) patron type (student, faculty, etc.); (6) type of transaction (i.e., check-in or checkout).³ From these archival tapes the number of circulations for the selected samples was extracted.

The Pearson correlation was calculated between the number of catalog card access points and the number of circulations within the sample size. The obtained r value was compared with the tabled r value within the constraints of degrees of freedom and significance level to reject or retain the null hypothesis.

Control of Contaminating Variables

The variance in library circulations can be accounted for by a number of sources: number of books available, number of users, number of copies available for a given title, number of courses offered, etc. In order to measure variability of selected variables (for this study, the number of card catalog access points for a given title and the number of times that title circulated in a given period), an effort was made to control as many contaminating variables as possible.

Since precise experimental control in social sciences research projects is not always possible, as many contaminating variables as possible were controlled through the random selection process. Sample data for this study were selected using a random number table.

One hundred titles from the education books (i.e., books that fall into the Library of Congress classification class L) acquired during the academic year 1976 and available in Booth Library as of August 16, 1977, were selected. Since Booth Library assigns an ac-

cession number to each physical volume acquired, a record of the range of numbers assigned on each working day is kept by the library, and since the number also becomes a data element in each machine-readable record representing the library's book collection, this selection was possible. Books made available for circulation on a given date can be easily identified with the aid of the accession number.

Circulation statistics of the 100 titles selected in the above manner for academic year 1977 and academic year 1978 were extracted. The above procedure ensured that the sample titles were available for circulation for an equal length of time and had equal chance of being checked out.

An obvious question arises as to whether newly acquired books might circulate more than older ones. A study by Bulick et al.⁴ found that the greatest use of books occurs during the first two years of their availability in a library. The Mueller study found that there was no significant difference in the circulation of new versus old books.⁵ Significant differences were found only when new books were displayed on a new-book shelf. Since Booth Library does not have a new-book shelf as such, an assumption can be made that the old and new books circulate alike.

In addition to the equal length of availability, the sample from academic year 1976 offers the following advantages:

1. Booth Library finished its reclassification project (i.e., switch from Dewey to LC classification) in the academic year 1976. As a part of this project, the card catalog had been completely revised and updated.

2. At the end of the reclassification project, the card catalog maintenance responsibility was placed in the hands of a single cataloger rather than a group of catalogers, as was the case before.

3. A systematic procedure of recording any changes in cataloging data for a given title on the back of its main entry card was initiated.

The third feature proved to be very useful for this study. It indicated that though subject headings did change from time to time, they did not increase the number of access points for a given title. The new subject heading simply replaced the old subject heading. Consequently, the number of access

points provided remained the same before and after the change of a subject heading.

The main reason for choosing education book samples is that during academic years 1977 and 1978 the School of Education enrollments deviated less than any other school in the university. The School of Education had an enrollment of 3,647 in fall 1977 and 3,755 in fall 1978.⁶ The number of students enrolled in a subject area contributes to the number of books circulated in that subject area. Selecting the years with less deviation in enrollments allows one to control any effects caused by the change in the number of students from year to year. Also, the number of courses offered by the School of Education remained the same during this period.

Additional reasons for choosing the education sample include a UCLA study which found that card catalog use was equally distributed between graduate and undergraduate students of education;⁷ and Shercliff's findings that 95 percent of the education book checkouts can be traced back to the education faculty and students. The latter study also found that though the education faculty and students check out books from other subject areas, the checkout of education books by students and faculty from other subject areas is negligible.

Sample Size

A sample size of 100 (about 10 percent of the education books acquired in the academic year 1976) was chosen based on the suggestion made by the statisticians Roscoe and McNeil. Roscoe⁸ states that a randomly selected sample of not fewer than 30, and not more than 500, is adequate for most behavioral research projects. Within the above limits, he suggests taking a sample of about a tenth of the population.

McNeil¹⁰ indicates that given a randomly selected sample of 100, one need not worry about Type II error since Type II error (i.e., the retention of a false hypothesis) is related to the sample size.

There appears to be no consensus among statisticians concerning sample size for a given problem. There are formulas to estimate the required sample size when the standard deviation of the variable being studied is known. Since that is not the case here, those formulas cannot be used.

Formulae

On the basis of experience and a review of the related literature, the following research method was used in this study:

1. $H: p = 0$ $A: p \neq 0$,
2. Pearson r , $N = 100$,
3. .05 level, two tailed, $df = 98$,
 $R: r \leq -.205$ or $r \geq .205$.

DATA COLLECTION

Booth Library maintains a machine-readable shelflist to support its online circulation system. There is a corresponding record in this machine-readable file for each physical volume available in the library collection. Each record in the machine-readable file contains the following data elements: author; title; call number; accession number; copy number; volume number; location code; part, supplement, or index number; and edition. With the aid of the accession number field, it is possible to determine on what day a particular item was placed in the library's collection. Consequently, a computer program was written to extract the data needed for this study.

In a series of steps, the computer was instructed to do the following:

Step 1: From the library's machine-readable shelflist, select all the titles belonging to the Library of Congress classification L, which is the class interval for the field of education. In this step, the computer selected 24,290 records representing the library's total collection of education books.

Step 2: From the records selected in step one, select all the records falling between the accession number range 343551-356283 (inclusive). The accession number range specified covers the academic year of 1976. In this step, the computer selected 827 records (considerably fewer than the estimated 1,000 volumes), representing the education books acquired during the academic year of 1976.

Step 3: From the records selected in step two, select all the records pertaining to the single-volume, single-copy books available within the circulating collection. In this step all multivolume and multicopy titles were eliminated along with the titles located in the reference collection. Inclusion of multivolume and multicopy titles will influence only one of the variables. Theoretically, the avail-

ability of multivolume and multicopy titles accounts for more circulation than the single-volume and single-copy titles. At the same time, multivolume and multicopy titles do not call for a corresponding increase of catalog access points in the public catalog. Consequently, they unduly influence one variable and may lead to biased results. For this reason, multivolume and multicopy titles were excluded from this study.

This step selected 462 records that represented the education titles acquired during the academic year of 1976 and met this study's criteria for selection. These 462 titles formed the population of this study.

SAMPLE SELECTION

One hundred random numbers falling within the range of the numbers 1-462 were selected from the random number tables in the Roscoe book, *Fundamental Research Statistics for the Behavioral Sciences* (2d ed., 1975).

When the selected random numbers were placed in numerical sequence, five duplicates were found. Consequently, an additional five numbers were selected. Thus a list of 100 unique random numbers was obtained. The range of the selected random numbers was 1-455. Using these random numbers, the corresponding titles from the alphabetical title list were selected. For example, the first random number was one and the second random number was four; consequently, titles occupying the first position and the fourth position in the list were selected.

For each title selected, the card catalog was consulted to determine the type and number of the catalog access points provided, and each entry was individually examined to be sure that it existed in the catalog. The pertinent information was recorded on the list across from each title selected. The recorded information was verified a few days later using the same process as before. Six errors were discovered and corrected during the verification process. Thus, the number of circulations and the number of catalog access points for each sample title was obtained.

DATA ANALYSIS

The numerical values representing catalog access points and the number of circulations

were input into a disk file by using a computer terminal. The data from the disk file was processed by the latest version (79.2B) of the Statistical Analysis System (SAS) available at the Eastern Illinois University Computer Center.

The PROC CORR procedure of the SAS was used to calculate the Pearson correlation. The PROC CORR procedure defaults to the Pearson correlation unless requested to do otherwise.¹¹

FINDINGS

1. The range of cumulative circulations for the sample studied is 0-21, with a mean of 3.59 and a standard deviation of 4.29.

2. The range of catalog access points for the sample studied is 2-7, with a mean of 4.28 and a standard deviation of 1.13.

3. Thirty-one percent of the titles selected have had a zero circulation rate during the study period. This approximately equals the library's previous finding that 33 1/3 percent of its circulating collection has no recorded circulation.¹² This indicates the representativeness of the random sample selected for this study.

4. The distribution of catalog access points for the 100 sample titles studied appears as follows:

No. of author (corporate or personal) main entries:	90
No. of title main entries:	10
No. of subject-added entries:	159
No. of added entries for editors, compilers, etc.:	12
No. of joint author-added entries:	34
No. of series-added entries:	10
No. of title-added entries:	90

In the above distribution, it is interesting to note that the average number of subject-added entries per title is 1.59, which is slightly higher than the national average of 1.3 subject-added entries per title as established by McClure in 1973.

CONCLUSION

The obtained r value is .136, which is less than the tabled r of .205. Consequently, the null hypothesis could not be rejected. On the basis of the data and the statistical analysis in this study, one should conclude that there is

no significant correlation between the number of card catalog access points per title and its recorded use in an academic library for a specified period of time.

RECOMMENDATIONS

1. It is recommended that this study be repeated at other academic institutions of different size, mission, and scope to determine whether these findings hold true at different institutions.

2. This study was primarily concerned with a linear relationship between the two variables and found there is no such relationship. However, there may be curvilinear relationships between these two variables, and it is suggested that such relationships be explored.

3. This study should be repeated, using titles in other subject fields as samples.

REFERENCES

1. R. C. McClure, "Subject and Added Entries as Access to Information," *Journal of Academic Librarianship* 2, no. 1:9 (1976).
2. *Ibid.*, p. 9-14.
3. P. V. Rao and B. J. Szerenyi, "Booth Library On-Line Circulation System (BLOC)," *Journal of Library Automation* 4:86-182 (1971).
4. S. Bulick and others, "Use of Library Materials in Terms of Age," *Journal of the American Society for Information Science* 27:175-82 (1976).
5. E. Mueller, "Are New Books Read More Than Old Ones?" *Library Quarterly* 35:166-72 (1965).
6. *Academic Information* (Charleston, Ill.: Eastern Illinois University, 1978).
7. *Library and User Requirements and Orientation Techniques* (Los Angeles: UCLA, 1974). ED 125 524.
8. W. H. Shercliff and others, *College of Education Libraries Research Project* (Manchester, England: Didsbury College of Education, 1973).
9. J. T. Roscoe, *Fundamental Research Statistics for the Behavioral Sciences* (2d ed.; New York: Holt, 1975).
10. K. A. McNeil, *Testing Research Hypotheses Using Multiple Linear Regressions* (Carbondale, Ill.: Southern Illinois Univ. Pr., 1975).
11. *SAS User's Guide* (Raleigh, N.C.: SAS Institute, 1979).
12. *Program Review Booth Library 1979* (Charleston, Ill.: Eastern Illinois University, 1979).