

CHARLES SAGE, JANET KLAAS,
HELEN H. SPALDING, AND TRACEY ROBINSON

A Queuing Study of Public Catalog Use

The authors conducted a six-week queuing study of public catalogs in the Iowa State University library system. Data gathered are analyzed primarily to determine if routinely gathered library statistics can validly be used to predict catalog usage, to discover the ratio between the usage of the card catalog and the serials catalog, and to pinpoint the time of peak card catalog usage in order to measure more closely the rate of use. This measurement, then, provides one factor in a simulation model that can be constructed to predict accurately the number of devices needed for an alternative catalog format.

FROM JANUARY 8 through February 18, 1979, the authors conducted a queuing study in which they examined the use of the Iowa State University Library's public catalogs. The study was sponsored by the library's Committee on the Public Catalog and Its Alternatives. This committee was appointed in November 1977 and was charged with studying the present public catalogs and possible catalog format alternatives in light of several challenges facing the library in the next decade. These challenges include managing limited funds, planning and completing a major building addition, considering the use of increasingly available technology, and weighing Library of Congress' decisions to close its card catalog and implement the second edition of the *Anglo-American Cataloguing Rules*.¹

It became evident that knowledge of the actual number of public-catalog users would aid the committee in making confident decisions about the future of ISU's public catalogs.

Few detailed studies had been published itemizing the quantity of catalog users at specific times. Several unique local factors made application of available findings difficult.

Charles Sage is lead systems analyst, Iowa State University Library; Janet Klaas is a government publications/reference librarian, Iowa State University Library; Helen H. Spalding is head of technical services, University of Missouri-Kansas City Library; and Tracey Robinson is a systems librarian, Harvard University Libraries.

Therefore, the Queuing Study Subcommittee, which included the authors, was formed in November 1978 to plan and implement the systematic observation of people using the library's public catalogs during the peak periods of library use. We wanted to determine the extent to which the public catalogs are used, the relationship of public-catalog use to circulation and door-count statistics, and the areas of the card catalog that are most heavily used. Previously at Iowa State, estimation of catalog use had been largely intuitive or based upon questionnaire-survey approaches, which are difficult to evaluate objectively. The Iowa State University Statistical Laboratory offered suggestions for the design of a statistically valid procedure.

The design of this procedure had to take into account the peculiarity that in the ISU Library no serials appear in the card catalog. Instead, serials are listed in a computer-produced book catalog that is distributed to locations throughout the library system. Staff members at public service areas had firsthand, though unmeasured, experience of heavy use of the serials catalog. With the queuing study, we hoped to quantify the use of the serials catalog in relation to the use of the card catalog. In evaluating alternative catalog formats, the Committee on the Public Catalog and Its Alternatives included the possibility of reintegrating serials and monographs in a central, machine-readable system. The existing machine-readable serials records on tapes could form the base of a new

computer-output microform (COM) or online catalog.

Also, we wanted to test the validity of using circulation and/or door-count statistics as substitutes for actual observation of the number of arrivals at the catalogs. The number of catalog users in a given period of time is one of the variables needed to construct a simulation model that could be used to determine the number of catalog devices (COM or online) needed to serve library clientele efficiently. To complete the model, device-operating times and various hypothetical numbers of devices need to be supplied in order to calculate waiting times.

METHODOLOGY

In an attempt to gather data representing typical catalog use during the quarter, the study was conducted for a six-week period during which no university holidays or exams occurred. Observations were made during the following hours of the library's peak occupancy: 10:00 a.m., 11:00 a.m., 1:00 p.m., 2:00 p.m., 3:00 p.m., 4:00 p.m., and 8:00 p.m., Sunday through Friday (excluding 10:00 a.m. and 11:00 a.m. Sunday and 8:00 p.m. Friday).

The card catalog in the main building contains 565,000 monographic titles described on 3,725,008 catalog cards. Author/title and LC subject heading cards are in two separate files. The physical area is made up of 1,500 square feet containing forty-one 72-drawer card catalog cabinets and eight tables varying in size from two-by-six-feet to three-by-twelve-feet. The annual serials catalog, in book format, contains bibliographic information for 30,000 titles, as well as indexes to these titles by corporate body and subject, and a cumulative supplement.

A total of twenty-nine catalog sites were selected for the study. The card catalog sites included the main library and six special locations (one branch library and five reading room catalogs). Serials catalog sites included five staffed locations in the main library, eleven locations in the main library stacks, and six locations in the branch library and reading rooms.

Ten-minute observation periods were randomly assigned so that each ten-minute period within the designated peak-use hours would be sampled approximately five to six times (once for each day of the week). During

each of the 234 observation periods, the number of patron arrivals and staff arrivals at the card catalog in the main library was recorded. To facilitate the counting of arrivals, the card catalog area was roped off with only one entrance/exit point. At the end of each observation period the location of each patron within the card catalog area was marked on a schematic drawing of the area.

The card catalogs and serials catalogs in the branch library and reading rooms were observed simultaneously during alternate main-library card catalog observation periods. The serials catalogs in staffed locations in the main library were observed during the same alternate periods. Following reciprocal alternate observation periods, one of the eleven serials catalog stack locations was observed. Selection of the site was randomly determined based on the assumption that a sample of the stack locations would adequately represent serials catalog use on the floors and tiers of the main library.

Taking into account that several observation periods were accidentally missed, we collected a total of 2,327 sample counts.

RESULTS

Scatter Diagrams and Pearson's Correlations of Card Catalog Use to Circulation and Exit Counts

Interpretation of the data for this portion of the study began with the hypothesis that a functional relationship existed between card catalog use and three routinely measured and accumulated statistical variables: exit count, books circulated, and number of persons checking out books. If these correlations did indeed exist, we wanted to show how the value of card catalog use could be predicted from data collected in the circulation department and at library exit points.

Scatter diagrams were prepared plotting card catalog use on the vertical axis and circulation and exit counts on the horizontal axis. We applied a common statistical procedure for fitting a line to a scatter diagram, the least-square regression, using the Statistical Package for the Social Sciences (SPSS).² In addition to producing a scatter diagram, the subprogram calculated Pearson's r , r^2 , Significance, Standard Error of Estimate (SEE), Intercept, and Slope.

Table 1, column 1 is the result of plotting

TABLE 1
SCATTER DIAGRAMS AND PEARSON'S CORRELATIONS DATA

	Scatter Diagram 1 Number of Books Charged per Hour vs. Number of Card Catalog Users per 10 Minutes	Scatter Diagram 2 Number of Persons Checking Out Books per Hour vs. Number of Card Catalog Users per 10 Minutes	Scatter Diagram 3 Number of Persons Exiting Library vs. Number of Card Catalog Users per 10 Minutes
Correlation (R)	.19203	.21655	.27460
Standard Error of Estimate	5.31212	5.28443	5.20479
Number of Plotted Values	224	224	224
R Squared	.03687	.04689	.07540
Intercept (A)	8.85120	8.35253	7.76930
Significance	.00196	.00055	.00002
Slope (B)	.02848	.09171	.00593
Missing Values	10	10	10

number of catalog users per hour and number of books charged per hour. Actually, it is difficult to make a visual observation when the absolute value of r is less than .3. In table 1, column 1, $r = .19203$, which denotes little or no relation between the two variables. The variance is explained by r^2 . Examining r^2 one sees that only 3.7 percent of the variation observed in the public catalog use are explained by books charged by hour.

The sample size is sufficiently large ("Central-limit theorem")³ to justify the assumption that the sampling distribution is normal. The standard error of estimate associated with predicting card catalog use is ± 5.31 persons per ten-minute period.

The second comparison of number of persons checking out books with those using the card catalog, as shown in table 1, column 2, indicates a stronger correlation ($r = .21655$) although it is still quite low. The explained variance is only 4.7 percent.

Comparison of the number of persons exiting the library with those using the card catalog is shown in table 1, column 3. Again, r and r^2 are low and indicate little relationship between the two variables.

The numbers of persons exiting the library, persons checking out books, and books circulated fluctuate over a wider range of frequencies than does the number of card catalog users. These fluctuations are reflected in table 1 and illustrate the low correlations.

To further examine the correlation between the variables tested within duplicate time periods, it was decided to offset the hours of comparison so that circulation and exit counts trailed card catalog usage counts.

The greatest increase of correlation occurred with a one-hour differential between number of books charged and number of users at card catalog. The value increases from .19203 (see table 1, column 1) to .47074. The one hour differential explains 22 percent dependence of one variable on the other. Even so, in the best of circumstances, the relationship of card catalog use to books circulated still leaves 78 percent of the variance unexplained.

Determining Arrival Rates by Isolating "Peak of Peaks"

The low correlation between card catalog use and circulation or door-count statistics implies that these statistics are questionable surrogates for catalog use; therefore, it was our opinion that arrival rates at the catalog should actually be observed.

Using the same arrival-rate data as above, we attempted to determine a mean arrival rate that ultimately could be used as a factor to calculate the number of devices for alternate forms of a catalog. Mean arrival rates, broken down by day and by time can be seen in tables 2 and 3. The analysis of variance reveals that the hourly means are significantly different from each other, as are the daily means. It seemed as if disproportionately low catalog use on Friday and Sunday might be responsible for the difference observed among the daily means, so a second analysis of variance was done in which only the Monday through Thursday data were examined. When the Friday and Sunday counts are suppressed, there is no significant difference among the daily means while the difference among the hourly means is accen-

TABLE 2
AVERAGE NUMBER OF ARRIVALS AT THE CARD CATALOG
PER TEN-MINUTE OBSERVATION PERIOD, BY DAY

	Monday	Tuesday	Wednesday	Thursday	Friday	Sunday
\bar{x} :	12.9	12.7	11.9	12.1	9.6	9.8
n:	42	41	40	42	36	30

TABLE 3
AVERAGE NUMBER OF ARRIVALS AT THE CARD CATALOG
PER TEN-MINUTE OBSERVATION PERIOD, BY HOUR

	10:00 a.m.	11:00 a.m.	1:00 p.m.	2:00 p.m.	3:00 p.m.	4:00 p.m.	8:00 p.m.
<i>Including Friday and Sunday</i>							
\bar{x} :	9.7	10.9	13.1	15.8	12.8	8.3	10.2
n:	30	30	36	35	35	36	29
<i>Excluding Friday and Sunday</i>							
\bar{x} :	9.7	10.8	14.2	18.0	14.0	8.8	11.3
n:	24	24	24	23	23	24	23

tuated. These analyses indicate that mean arrival rates remain fairly constant across the four days, Monday through Thursday; however, within each day, arrival rates fluctuate from hour to hour.

In order to determine the *maximum* arrival rate, it was necessary to isolate the busiest times of use at our present catalog. To identify these peaks of all peak-use times, count frequencies were tabulated for the Monday through Thursday data. Observation periods which had an arrival rate of 15.5 or more (this constitutes less than 10 percent of the cases) were identified as shown on figure 1. Despite some scattering of peak-use times, which is to be expected given the small sample size (i.e., each ten-minute period was observed only once for each day of the week), the heavy-use periods cluster between 2:00 and 2:30 p.m. across all four days. We thus hypothesized that the mean arrival rate between 2:00 and 2:30, Monday through Thursday, is relatively constant.

The above time blocks were tested to determine the goodness of fit of the data to certain standard distributions. We had too little data for the 2:00 to 2:30 p.m., Monday through Thursday period to make that determination. Additional data were collected from January 8 through February 18, 1980 (one year after the original study, during a comparable six-week period). Observation peri-

ods were limited to the peak periods of 2:00 to 2:30 p.m., Monday through Thursday. Counts were recorded for each minute, rather than for ten-minute periods, in order to increase the precision of the measure of arrival rate. A total of 690 one-minute samples were recorded. Table 4 shows the average number of arrivals per minute for each of the thirty one-minute observation periods with a mean arrival rate of 1.9 persons per minute over the entire sample.

With some reliance on the literature,^{4,5,6} we compared the distribution pattern of arrivals with a Poisson distribution using the chi-square goodness-of-fit test. In most human situations, arrival rates are not fixed so that there would be a steady, even number of users every minute. The Poisson distribution can be used to predict queueing, given a range of random arrival rates with no dependence upon each other, and an average arrival rate. (The mean arrival rate, Monday through Thursday, is relatively constant.) The results in table 5 show that at the 95 percent confidence level, the observed data follow a Poisson distribution with a predicted arrival rate of 1.9 persons per minute. There is a high probability that one to three people will arrive each minute at the card catalog during peak use periods. This information, combined with the average use time of a catalog reader or terminal and tolerable waiting times dur-

Frequency of Arrivals at the Card Catalog by Day and Time (per Ten-Minute Sample Period).

Fig. 1

(The dotted line at 15.5 isolates the top 10 percent, or the peak of the peak-use periods.)

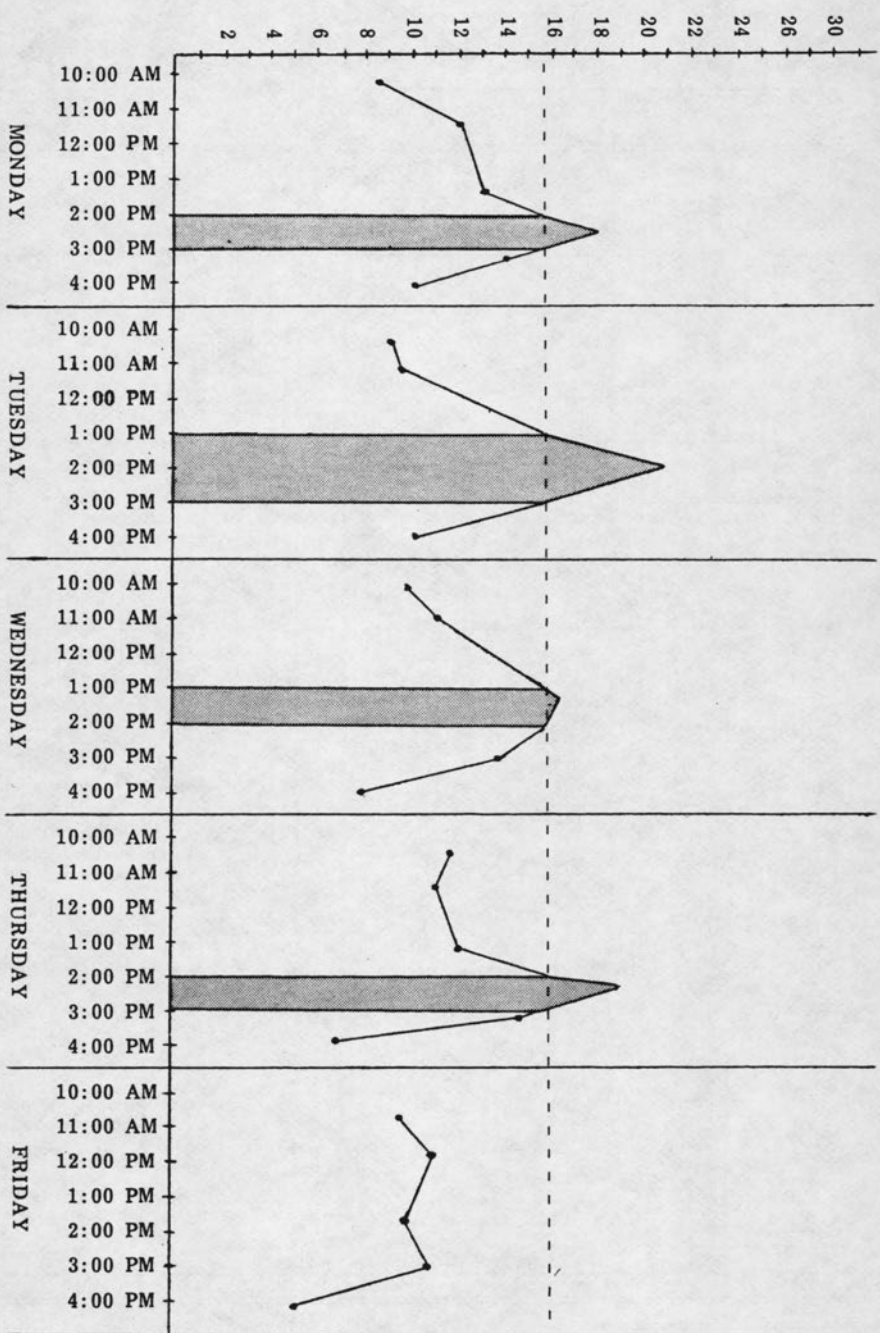


TABLE 4
AVERAGE NUMBER OF ARRIVALS (PUBLIC ONLY) AT THE
CARD CATALOG, PER MINUTE, 2:00-2:30 P.M.

Week	Monday	Tuesday	Wednesday	Thursday	Totals	
1	Mean	1.73	1.8	1.77	1.77	
	Number counts taken	30	30	30	90	
	Number users counted	52	54	53	159	
	Standard Deviation	.87	1.37	1.36	1.21	
2	Mean	1.67	1.3	1.56	2.1	1.66
	Number counts taken	30	30	30	30	120
	Number users counted	50	39	47	63	199
	Standard Deviation	1.37	1.21	1.25	1.49	1.39
3	Mean	2.13	1.5	1.35	1.04	1.31
	Number counts taken	30	30	30	30	120
	Number users counted	64	45	57	46	212
	Standard Deviation	1.57	1.2	1.35	1.04	1.31
4	Mean	2.03	2.23	2.37	1.8	2.11
	Number counts taken	30	30	30	30	120
	Number users counted	61	67	71	54	253
	Standard Deviation	1.27	1.43	1.25	1.21	1.3
5	Mean	2.47	2.67	2.43	2.0	2.34
	Number counts taken	30	30	30	30	120
	Number users counted	74	74	73	60	281
	Standard Deviation	1.38	1.93	1.63	1.49	1.61
6	Mean	2.4	1.33	1.83	1.3	1.72
	Number counts taken	30	30	30	30	120
	Number users counted	72	40	55	39	206
	Standard Deviation	1.8	1.4	1.37	1.06	1.48
Totals	2.14	1.76	1.98	1.75	1.9	
	150	180	180	180	690	
	321	317	357	315	1,310	
	1.5	1.43	1.42	1.3	1.4	

TABLE 5
COMPARISON OF THE OBSERVED DISTRIBUTION OF ARRIVALS AT THE
CARD CATALOG WITH THE EXPECTED DISTRIBUTION, ASSUMING A
POISSON DISTRIBUTION AND AN AVERAGE OF 1.9 ARRIVALS PER MINUTE

Number of Arrivals (x)	Probability f(x)	Expected Number of One-Minute Periods with x Arrivals	Observed Number of One-Minute Periods with x Arrivals	Chi-Square, Goodness-of-Fit-Test
0	.15	103.5	108	.196
1	.285	196.6	195	.013
2	.27	186.3	185	.009
3	.17	117.3	113	.158
4	.08	55.2	54	.026
5	.03	20.7	24	.526
6 (or more)	.014	9.7	11	.254
Total	.999	(689.3)	690	1.182

TABLE 6A
SERIALS BOOK CATALOG/CARD CATALOG COMPARISON:
SERIALS BOOK CATALOG USE

Location (Number of Copies)	Average Number of Users	Number of Observation Periods	Standard Deviation
Reserve (1)	.50	117	.71
Circulation (2)	2.66	115	1.50
Reference (4)	5.57	116	3.36
Periodicals (2)	4.74	117	2.67
Government Publications (1)	.21	115	.41
Floors/Tiers (11)	3.88	119	6.60
Totals	17.56	699	8.06

TABLE 6B
SERIALS BOOK CATALOG: CARD CATALOG: COMPARISON:
SERIALS BOOK CATALOG: USE AND CARD CATALOG: USE

Location	Average Number of Users	Number of Observation Periods	Standard Deviation (S or $\sqrt{\sum S^2}$)	Standard Error	SE/ \bar{x}
				$\frac{\sum s^2}{n}$	
Card Catalog	11.61	231	5.40	.36	.03
Serials Book Catalog	17.56	699	8.06	.30	.04

ing peak times, can be used to predict the number of devices needed to serve catalog users efficiently.

Serials Book Catalog/Card Catalog: Use Comparison

The distribution of multiple copies of the serials book catalog throughout the library complicated the process of comparing the frequency of its use to the frequency of card catalog use. The comparison was based on the assumption that the many copies of the serials book catalog represent only one intellectual body of information. By combining the use frequencies for all six locations, the total, frequency for serials catalog use was obtained (table 6A).

As previously described, five of the locations observed are staffed locations. The sixth location is an amalgamation of eleven copies of the serials catalog distributed throughout the library stacks. We assumed the frequency of serials catalog use at any of these stack locations would be representative of all eleven locations. Each of the eleven copies was observed approximately eleven times, thus providing 119 samples.

Patron use of the card catalog and total serials catalog use (at sixteen locations, accounting for twenty-one copies of the serials catalog) are shown in table 6B. Again, assuming the sampling distribution is approximately normal, the 95 percent confidence interval for card catalog use suggests that the average frequency of use falls within the range of 10.9 to 11.6 users per ten-minute period. The comparable confidence interval for serials catalog use is 16.95 to 18.2 users per ten-minute period.

The ratio of serials catalog use to card catalog use (calculated from the means) is approximately 60:40. By taking the confidence intervals into account, this ratio ranges between 63:37 to 58:42. These results indicate that the serials catalog is used extensively and

is probably consulted more frequently than the card catalog by patrons, especially in light of the fact that not all copies of the serials catalog were observed in this study.

Placement of Users within Card Catalog Area

Measurements of the placements of card catalog users within the catalog area were accomplished by marking the location of all people inside the roped-off area at the end of each of the 234 ten-minute counting periods.

Though admittedly employing a crude means of measuring catalog-use distribution, we nevertheless extracted from the data some useful information. We counted a total of 2,206 people within the area (an average of 9.4 per count). Of these, 1,136 (51 percent) were located in the author/title section of the area and 1,070 (49 percent) in the subject section as shown in Table 7.

Several surveys have been done which examine the types of catalog searches performed by users. Lipetz found that about three-fourths of all searches are (at least initially) known-item (i.e., document or author) searches.⁷ Tagliocozzo and Kochen found that

TABLE 7
CARD CATALOG USE BY SECTIONS

Section	Users
<i>Author/Title Section</i>	
A-Brown, R (3 units)	162
Brown, S-DIGF (3 units)	144
DIGG-IMPERIAK (5 units)	227
Imperial-New York, Unit (5 units)	191
New York, Univ-Steep (5 units)	215
Steep-ZZ (4 units)	197
Total	1,136
<i>Subject Section</i>	
A-CIU (3 units)	201
CIV-FRID (3 units)	220
FRIE-MASS (3 units)	216
MASSA-Rail (3 units)	210
RAIM-Z (4 units)	223
Total	1,070

approximately 65 percent of the searches analyzed were known-item searches.⁸ It is difficult to compare these findings to the results of this study in light of the different methodology used (i.e., observed user locations as opposed to user interviews). Assuming that the use of a catalog (subject or author/title) provides a rough indication of the type of search being performed, a tentative comparison can be made.

Both of these studies previously mentioned, based upon actual interviews with catalog users, had the advantage of user-status data. From these data, the investigators could see that use of the subject searches declined as the status of the users increased. Underclasspersons tended to perform subject searches proportionately more than upperclasspersons who, in turn, used subject searches more than graduate students and faculty. As expertise grows in a field, the researcher tends to use the more specialized indexes and abstracts to journals and nonconventional report literature and to be familiar with specific authors and titles. Conversely, undergraduates are the group most likely to be searching the subject catalog.

This observation helps to explain the relatively low (51 percent) use of the author/title catalog observed in this study. Iowa State University's graduate and research programs are concentrated in technology and the sciences—areas in which literature searches are drawn principally from journal and report literature. A related factor affecting the use of the author/title catalog at Iowa State is the autonomous serials catalog. A person scanning the indexing and abstracting services or using a computer-generated subject bibliography and locating a particular journal title of interest does not then search for that title in the card catalog (a typical known-document search), but instead locates the title in the serials book catalog. Comparison of serials catalog and card catalog use at ISU, indeed, shows heavier use of the serials catalog. We can, therefore, speculate that the use of the serials catalog lowers the use of the author/title section of the card catalog because of the elimination of serials title searches.

Table 7 shows a more specific breakdown of the distribution of users at the card catalog. Use is fairly evenly distributed throughout

the catalog area, suggesting that no one specific area is used more heavily than any other.

CONCLUSION

The purpose of this study was to collect information and data that would provide a description of catalog use at the Iowa State University Library. Special emphasis was placed on accurate assimilation of actual counts, and we avoided as much as possible any subjective analysis. Every precaution was taken to preserve the statistical integrity of the results, and confidence factors were calculated wherever feasible. The chronology of our study had to be altered when we discovered we did not have enough data to accurately calculate the distribution of arrivals at the card catalog during peak-use times after we had identified those times; however, the additional observations of catalog arrival rates one year later enabled us to verify the assumption that the Poisson distribution provides an accurate portrayal of the distribution of arrivals.

Low correlations between regularly gathered statistics (i.e., circulation statistics and door counts) and catalog arrival rates suggest rather strongly that these statistics are not good surrogates to predict catalog arrival rates and use. As mentioned in the introduction, we did not measure patron service times for primarily two reasons. The first reason was that we could not devise a scheme that was economically practical to implement. We either needed a dozen or more observers or an elaborate video camera scanning arrangement, neither of which we could afford. Second, the determination of card and book catalog service times would provide very little utility. We have seldom had waiting lines form at either the card catalog or at the book catalogs, so under current arrangements little could be done to improve physical access. To use these service times as substitutes for automated-device service times in a queueing model would be in error if one were trying to calculate waiting-time distributions or device-quantity requirements. To complete our queueing model we expect to devise hypothetical service times based on actual on-line catalogs or COM catalogs operating at other installations.

Another objective of this study was to compare monographic catalog and serials catalog use. The physical separation of these catalogs,

with basically identical access points, gave us a unique opportunity for comparison. The high use of the serials book catalogs (60 percent of all catalog arrivals) points out that a relatively small percentage (6.6 percent) of our total collection of titles is being accessed heavily. This fact, coupled with the fact that active serials experience frequent record changes, indicates that special priority should be given serials when converting to machine-readable form for online access. A less significant but interesting factor is the potential impact this breakdown of usage will have on acquisition decisions and technical processing priorities within the Iowa State University Library.

Initially we had hoped that a delineation of the exact positions of users within the card catalog area would be helpful in the event that the ISU Library administration chose a COM format for our future catalog. Since the use of our divided author/title-subject card catalog was almost evenly split, and because distribution of users throughout the entire catalog area was fairly even, there are two options for the design of a microfilm catalog. The catalog, in film form, could remain di-

vided with an equal number of readers for the author/title and the subject sections; or a dictionary catalog could be produced with each reel containing a set number of records (analogous to the sections of the card catalog containing a set number of cards). These two plans, however, are only feasible with a separate serials catalog.

Two components of our study that are omitted from this paper are a patron-staff comparison of the card catalog use, and a delineation of branch library and reading room card catalog use. This information, we feel, is largely of parochial interest but is available, upon request, from the authors.

The overall results of this study produced a few surprises, but we are satisfied these analyses are valid and will help us make confident decisions concerning the future of bibliographic access to our collection. Obviously certain peculiarities of the catalog system of the Iowa State University Library affected our findings, but application of our methodology should prove helpful to other libraries whose staffs are adapting their catalogs with the new technology.

REFERENCES

1. Michael Gorman and Paul W. Winkler, eds., *Anglo-American Cataloguing Rules* (2d ed.; Chicago: American Library Assn., 1978).
2. Norman H. Nie, *SPSS: Statistical Package for the Social Sciences* (2d ed.; New York: McGraw-Hill, 1975).
3. Hubert M. Blalock, Jr., *Social Statistics* (New York: McGraw-Hill, 1979).
4. Abraham Bookstein, "Congestion at Card and Book Catalogs: A Queuing Theory Approach," *Library Quarterly* 42:316-28 (July 1972).
5. Ronald W. Force and Jo Ellen Force, "Access to Alternative Catalogs: A Simulation Model," *College & Research Libraries* 31:234-39 (May 1979).
6. Philip M. Morse, *Library Effectiveness: A Systems Approach* (Cambridge, Mass.: MIT Press, 1968).
7. Ben-Ami Lipetz, "Catalog Use in a Large Research Library," *Library Quarterly* 42:129-39 (Jan. 1972).
8. R. Tagliacozzo and M. Kochen, "Information-Seeking Behavior of Catalog Users," *Information Storage and Retrieval* 6:363-81 (Dec. 1970).