

Research Notes

Observations of Browsing Behavior in an Academic Library

Johanna Ross

The increasing costs of land and buildings, coupled with inflationary pressures on operating budgets, are forcing library administrators to consider nontraditional methods in providing patron access to collections. Instead of building traditional library space, as buildings become crowded, warehouse-type structures are being constructed on less-expensive land. Items with low circulation records are being stored on compact shelves to be paged as requested. Thus, traditional browsing access to large, open shelf library collections may be substantially curtailed or even eliminated. This is particularly true at the Physical Sciences Library at the University of California at Davis, where one volume will soon need to be moved off site for each new volume added.

Volumes to be moved off site must be selected in a manner that minimizes inconvenience and provides for maximum document retrieval. The circulation record is an adequate measure of usage for those items actually borrowed, but it does not take into account information gathered in-house, i.e. those items that are used at the shelf and/or left for refiling.

Browsing activity, or unrecorded use as it is sometimes called, has been of concern to librarians since the advent of the open shelf policy, but arriving at estimates of this activity has been most difficult. Some librarians have attempted to overcome

this problem by counting refiles as browsing use. Although this may provide clues to the noncirculating use of the collection, refiles measure only a portion of in-house use. Taken alone, counting refiles is not an adequate indication of browsing activity. Shelf consultations do not result in refiles or circulation, where statistics can be measured, compared, collated, or classified. Patrons are free to remove books in an open stack situation, consult them briefly or at length, and then replace them. Only the patron is aware that the item was used and only he can assess the value of the information acquired.

Browsing research efforts have generally focused on two general techniques. The first consists of surveying patrons entering or leaving the stack area. Questionnaires designed to measure the nature and extent of the browsing activity are given to the patron to be filled out.¹ Certain methodological problems are associated with this technique. If patrons fill out information after the fact, knowledge that their activity is being recorded may influence memory. Then there is the problem of the definition of shelf consultation. Is it considered a shelf consultation only when information is actually found, or does consulting the table of contents or index of a book constitute a shelf consultation? If some patrons consider this trivial information not worthy of reporting, while others

Johanna Ross is associate librarian, Physical Sciences Library, University of California at Davis.

do not, responses would not be considered reliable.

The second technique, pioneered in the work described by Simon and Fussler,² relies on questionnaires inserted into books that have been statistically selected. With this method it was possible to determine that books had been disturbed, even if the included questionnaires were not filled out. Pens were provided for a portion of the sample as an incentive tool and the patron was instructed to keep the pen after filling out the questionnaire. A significantly greater number of the questionnaires that had pens attached were filled out than those that did not.

A wide range of shelf consultation has been reported in the literature. Slater and Fisher estimated that the average number of items consulted were 4.1 and the average found useful were 2.4.³ The study by Simon and Fussler estimated that three to nine books were consulted for every recorded use.⁴ C. Harris,⁵ in an analysis of refiles, together with questionnaires inserted in books statistically selected, estimated that as many as twenty times the number of books are used in-house than are checked out.

With a low figure of three consultations per recorded use and a high of twenty, these research results reveal a wide range in estimates of browsing activity. Some variation may be accounted for by the different subject content of collections and different regulations concerning stack use. What have not been considered as a source of variation are the different interpretations of the meaning of the word *browse*. Does it mean extensive use at the shelf or at a study area? Does it mean a brief consultation? Different concepts of the meaning of the word by those filling out the questionnaires may account for some of these variations.

If patrons had been actually observed, the word could have been standardized for the entire study. But researchers have not yet attempted to assess browsing activity in this manner. Upon examination, the methodological difficulties in assessing browsing use by observation did not appear any more difficult than doing it by inserting questionnaires into books statis-

tically selected. It was decided to undertake such a study where browsing would be defined as the examination of books at the shelf. Browsing intensity would be measured by recording the number of books removed *and* replaced and the time spent performing the activity. Those items that were taken from the area were not included as browsing, since these statistics would be recorded either as circulations or refiles. Although circulations and refiles were not considered as part of browsing activity, both were recorded and are given for comparison purposes (see appendix A).

To be able to generalize about the methodology, observed patron activity must be selected at random. There are two aspects to the sample selection process—time and place. To select time, the hours of opening were divided into fifteen-minute intervals. Random tables were used to select the time interval, then to select the day of the week. To select starting places, the non-reference shelf ranges were numbered and shelves were selected at random for each fifteen-minute interval.

Forty fifteen-minute intervals were observed for each week of the study. Thirteen weeks, one entire quarter and inter-session, were sampled in order that observed usage patterns would be representative and not influenced by a specific time during the instruction period.

Since only one patron could be observed at a time and many individuals could be engaged in browsing activity, the entire library was inspected five minutes before the selected time interval. The place of browsing and time of day were recorded for every patron observed browsing. When this was finished, the first patron located after the randomly selected starting point had browsing behavior recorded. If it was the same person seen on the initial walk-through, the patron was not recounted.

When a patron was observed standing at the shelf examining books, an inconspicuous place providing full view of browsing activity was chosen. Using a clipboard with data forms attached, the observer would pretend to consult the shelves for needed information. To the pa-

tron it would appear that the library staff member was gathering independent information. In reality, patron activity at the shelf was being recorded. This unobtrusive technique appears to have worked well. To our knowledge, none of the patrons were aware that their browsing behavior was the object of study.

The day of the week and time of day, as determined by random selection, were already recorded on the form, together with starting location. The browsing area, given in LC class notation, time duration, and the number of books removed and replaced were the remaining items recorded for each observation. Net browsing was defined as the number of books replaced. If a patron left the book-stack area with book in hand, it was assumed that the book would be counted either as circulation or refile.

RESULTS

Of the 520 fifteen-minute intervals selected, there were only 98 time periods when no one was browsing in the library at the selected time. In the remaining 422 sessions, 515 patrons were observed; 384 were first persons seen, while 131 observations were the second or third persons seen during selected intervals.

Circulation and refile counts were kept for the same time period to be compared with browsing data. Refiles are not included in the estimate of browsing use. Refiles are those items removed from the shelf but left on tables, in refile bins, or on refile shelves. Since circulations and refiles are actual counts and browse figures are based on a sample, the figures for all categories were first categorized by LC class number and then the numbers were reduced to percentages for meaningful comparisons. Thus, observed differences in use for the various classes can be seen more readily (see appendix A).

Books Removed

The number of books that patrons removed clusters at the low end of the distribution. Just over 50 percent of the patrons removed two books or less. Although the percentage of those who removed more than two books declines steadily, the

number of books they removed rises. Nearly 25 percent inspected five or more books and approximately 4 percent removed ten or more (see appendix B).

R^2 , or the coefficient of determination, indicates that time alone does not account for the number of items removed ($R = .339$; $R^2 = .1149$). Only 11 percent (R^2) of the total variation was explained by this relationship (see appendix C).

Sample means were used in this study as point estimates of central tendency. However, because these estimates are based on survey sample data, the figures are subject to potential sampling error. Additional estimates commonly used to indicate the extent of statistical reliability of sample data are interval estimates or "confidence intervals."

Confidence intervals specify ranges within which we can expect the actual value to be, with a certain level of confidence. The confidence interval (CI) of the sample means, \bar{X} , as an estimate of the actual mean, can be expressed as: $CI = \bar{X} \pm ts/\sqrt{n}$; where \bar{X} = the same mean; t = the confidence factor associated with the particular level of confidence desired; for the 95 percent confidence level, $t = 1.96$; s = the standard deviation of sample observations about the mean; n = the sample size; s/\sqrt{n} = the standard error of the mean.

In this study the observed mean number of books removed was 3.37 and the standard deviation was 2.72 (see appendix D). Since it was assumed that patrons were observed halfway through the browse, the mean number of books removed, \bar{X} , increases to 6.74 and the standard deviation to 5.44. For the 95 percent confidence level ($t = 1.06$), these figures applied to the above formula with 495 sample observations (n): $CI = 6.74 \pm 1.96 (5.44)/\sqrt{(495)} = 6.74 \pm .48$.

At the 95 percent confidence level, the true number of books removed per patron was between 6.26 and 7.2 ($6.74 \pm .48$).

Books Replaced

The mean number of books replaced was 2.26 (see appendix E). Time spent replacing books had little to do with the number replaced. R^2 , or the coefficient of

determination, revealed that only 13 percent of the total variation was explained by this relationship ($R^2 = .1289$; see appendix F). The number of books replaced correlates strongly with the number of books removed ($R = .907$; see appendix G).

Again, applying the confidence interval we find: $CI = 5.52 \pm 1.96 (5.0)/\sqrt{(495)} = 5.52 \pm .48$.

In short, at the 95 percent confidence level, the true number of books replaced per patron was between 5.02 and 6.0 ($5.52 \pm .48$).

Time Duration

While the mean time patrons spent at the shelf was 6.94 minutes (this becomes 13.88 when doubled because of being observed midway through the browse), approximately one-third of the browsing sessions were for three minutes or less (see appendixes H and I). Most of the 138 patrons who removed only one book and the 8 patrons who never removed a single volume, but simply scanned spine titles, did not engage in extended browsing (appendix B). This may be an indication that these users were simply looking for a book by call number and may not have had browsing as the principal reason for coming to the shelf. Once at the shelf, however, patrons usually examine more than one book before leaving, with or without one or more. A frequency time table is provided in appendix J.

After the first five minutes, the percentage of patrons leaving begins to decline. If the dips at four, six, and fourteen minutes are "seen" as a human tendency to round off time (a stopwatch was not used since observers either used their own watch or

the clock on the wall), the percent of decrease in the number of patrons with the increase of time appears smoother. Approximately 20 percent of browsing activity is for fifteen minutes or more (or assuming they were caught at midpoint, thirty minutes). Though no attempt to qualify browsing was made, some patrons were obviously very engrossed in the activity (see appendix I). $CI = 13.88 \pm 1.96 (10.06)/\sqrt{(486)} = 13.88 \pm .89$.

At the 95 percent confidence level, the time spent at the shelf is between 12.99 and 14.77 minutes.

Summary and Conclusion

This study, based on observed browsing behavior, corresponds to the findings of the Simon and Fussler study. At the 95 percent confidence level, the number of books replaced per patron was between 5.02 and 6.0. This narrows the range of 3 to 9 books they reported. This estimate falls short of the estimate of 20 made by Harris, but the Harris study counted refiles as part of browsing.

This study has employed the unobtrusive observation technique to determine browsing activity. The technique is simple to use and can be replicated with only a modest budget. If the survey conducted at the Physical Sciences Library at the University of California at Davis were replicated and extended to other university libraries and different subject collections, a broader range of browsing behavior could be accumulated. This would enable librarians to have a firmer foundation on which to base decisions regarding the management and access to materials shelved in off-site locations.

REFERENCES

1. The most thorough survey was done by Slater and Fisher when they surveyed 600 libraries; 104 cooperated and 6,300 usable questionnaires were analyzed. Margaret Slater and Pamela Fisher, *Use Made of Technical Libraries* (Aslib Occasional Pub., 2 [London: Aslib, 1969]), p.3.
2. Herman H. Fussler and Julian L. Simon, *Patterns in the Use of Books in Large Research Libraries* (Chicago: Univ. of Chicago Pr., 1961).
3. Slater and Fisher, *Use Made*, p.3.
4. Fussler and Simon, *Patterns in the Use*, p.115.
5. C. Harris, "A Comparison of Issues and In-Library Use of Books," *Aslib Proceedings* 20:118-25 (Mar. 1977).

APPENDIX A

LC Class	Total Books	Total Circ	Total Refile	Observed Browsers	Total Browsers	Books Removed	Books Replaced	Total Library Collection Percent	Circulation Percent	Refile Percent	Browsing Percent	Book Removed Percent	Books Replaced Percent
A-F	112	5	11	0	5	0	0	.08	.04	.13	.45	.00	.00
G-GZ	1,279	185	69	1	4	6	6	.96	1.55	.84	.36	.37	.56
H-P	1,404	146	91	6	12	17	12	1.06	1.22	1.11	1.08	1.06	1.12
Q	3,784	267	209	10	29	32	15	2.85	2.23	2.55	2.61	2.00	1.39
QA	16,769	2,069	1,316	78	166	302	221	12.64	17.30	16.03	14.94	18.86	20.54
QB	2,723	170	93	8	19	25	14	2.05	1.42	1.13	1.71	1.56	1.30
QC	18,522	1,519	950	60	140	205	134	13.96	12.70	11.57	12.61	12.80	12.45
QD	19,267	1,685	1,599	89	209	276	174	14.52	14.09	19.47	18.81	17.24	16.17
QE	13,839	1,260	398	27	74	99	56	10.43	10.54	4.85	6.66	6.18	5.20
QF-QZ	1,375	175	107	6	16	25	15	1.04	1.46	1.30	1.44	1.56	1.39
R	254	29	24	0	1	0	0	.19	.24	.29	.09	.00	.00
S	802	136	61	2	5	5	2	.60	1.14	.74	.45	.31	.19
T	3,693	206	97	3	12	11	4	2.78	1.72	1.18	1.08	.69	.37
TA	8,963	856	568	42	94	146	102	6.75	7.16	6.92	8.46	9.12	9.48
TC	2,667	222	165	16	31	67	46	2.01	1.86	2.01	2.79	4.19	4.27
TD	3,048	460	226	16	30	63	50	2.30	3.85	2.75	2.70	3.94	4.65
TE	710	78	23	0	7	0	0	.53	.65	.28	.63	.00	.00
TF	68	4	4	0	1	0	0	.05	.03	.05	.09	.00	.00
TG	167	6	8	0	1	0	0	.13	.05	.10	.09	.00	.00
TH	1,217	148	99	2	6	3	2	.92	1.24	1.21	.54	.19	.19
TJ	4,069	364	234	17	34	60	38	3.07	3.04	2.85	3.06	3.75	3.53
TK	7,456	759	746	26	78	72	49	5.62	6.35	9.08	7.02	4.50	4.55
TL	4,829	466	313	24	46	73	46	3.64	3.90	3.81	4.14	4.56	4.28
TN	5,234	331	201	17	37	66	58	3.94	2.77	2.45	3.33	4.12	5.39
TP	5,444	326	428	20	44	41	27	4.10	2.72	5.21	3.96	2.56	2.51
TR	520	30	29	1	3	1	0	.39	.25	.35	.27	.06	.00
TS	1,253	35	68	2	3	3	2	.94	.29	.83	.27	.19	.19
TT	11	1	1	0	0	0	0	.01	.01	.01	.18	.00	.00
TX-VM	189	9	19	1	2	3	3	.14	.07	.23	.18	.19	.28
Z	3,055	13	55	0	2	0	0	2.30	.11	.67	.00	.00	.00
	132,723	11,960	8,212	474*	1,111	1,601*	1,076*	100.00	100.00	100.00	100.00	100.00	100.00

*Based on sample only.

Note: There were 41 cases where LC classification was not recorded.

APPENDIX B

Books	Removed Number of Cases	Percentage	Replaced Number of Cases	Percentage
	(n)		(n)	
0	8	1.59	130	26.26
1	138	27.44	111	22.43
2	113	22.46	91	18.38
3	68	13.52	60	12.12
4	53	10.54	26	.25
5	40	7.95	29	5.86
6	24	4.77	14	2.83
7	16	3.18	8	1.62
8	14	2.78	8	1.62
9	10	1.99	5	1.01
10	5	.99	5	1.01
11	7	1.39	4	.81
12	3	.60	2	.40
13	1	.20	1	.20
14	1	.20	0	0
15	2	.40	1	.20
16	0	0	0	0
17	1	.20	1	.20
18	0	0	0	0
19	0	0	0	0
20	0	0	0	0
21	1	.20	0	0
	503*	100	495†	100

* 9 missing cases.

†17 missing cases.

APPENDIX C

Chi Square = 246.37583 with 132 Degrees of Freedom

Cramer's V = 0.21715

Contingency Coefficient = 0.58441

Pearson's R = 0.33933

APPENDIX D

NUMBER OF BOOKS REMOVED

Variable V8		Std. Error		Std. Dev.	
Mean	3.366	0.123		2.72	
Variance	7.431	1.000		Maximum	21.00
Range	20.000				
Sum	1666.000				
Valid Observations =	495				

APPENDIX E
NUMBER OF BOOKS REPLACED

Variable V9					
Mean	2.26	Std. Error	0.131	Std. Dev.	2.496
Variance	6.232	Minimum	1.000	Maximum	17.000
Range	16.000				
Sum	1119.000				
Valid Observations = 495					

APPENDIX F

V7 Time Duration by Number of Books Replaced
 Chi Square = 203.42319
 Cramer's V = 0.22986
 Contingency Coefficient = 0.60628
 Pearson's R = 0.35935

APPENDIX G

V7 Number of Books Removed
 by
 V9 Number of Books Replaced
 Chi Square = 1634.47894
 Cramer's V = 0.63804
 Contingency Coefficient = 0.90413
 Pearson's R = 0.90698

APPENDIX H

Variable V7		Time Duration			
Mean	6.944	Std. Error	0.228	Std. Dev.	5.031
Variance	25.306	Minimum	1.000	Maximum	20.000
Range	19.000				
Sum	3375.000				
Valid Observations 486					

APPENDIX I

TIME DURATION

Minutes	Number of Cases	Percentages
1	50	10.29
2	57	11.73
3	54	11.11
4	34	7.00
5	73	15.02
6	11	2.27
7	26	5.35
8	24	4.94
9	15	3.09
10	19	3.91
11	12	2.47
12	9	1.85
13	9	1.85
14	1	.20
15	83	17.08
16	1	.20
17	1	.20
18	3	.62
19	0	0
20	4	.82
Total	486	
Missing Data	29	

APPENDIX J

BROWSING FREQUENCY BY TIME OF THE DAY*

Time	Absolute Frequency	Browsing Rel. Freq. (percent)	Incidence of No Browsing
8-9 a.m.	25	2.1	15
9-10	76	6.3	6
10-11	122	10.0	9
11-12	128	10.5	3
12-1 p.m.	73	6.0	6
1-2	136	11.2	0
2-3	92	7.6	6
3-4	136	11.2	3
4-5	147	12.1	1
5-6	84	6.9	8
6-7	32	2.6	12
7-8	50	4.1	13
8-9	41	3.4	8
9-10	43	3.5	6
10-11	29	2.4	5
Missing	1	.1	0
Total	1,215	100.00	101

*The figures given include patrons encountered on the initial walk-through that did not have browsing behavior recorded. The time of day, day of the week, and place of browsing are recorded for 703 patrons. Only 1 patron could be observed at a time so full data is available for 512 individual browsing sessions. There were a total of 1,215 instances of browsing during the time intervals chosen.