

Measuring Collections Use at Virginia Tech

Paul Metz and Charles A. Litchfield

This study of the Virginia Tech Library collections brings together data on circulation and in-house use, including that of current periodicals, to assess differences according to kinds of use, variations in measurement technique, and time period. Use of current periodicals is found to be both large and qualitatively different from other kinds of use. Differences between circulation and in-house use are discussed. Despite these differences, circulation and in-house use are generally similar in their subject distribution. The subject distribution of circulation patterns is remarkably stable over time, and can be reliably assessed using short sampling periods.



ver the past several decades, many authors have contributed to our understanding of how library collections are used. They have identified the characteristics of books receiving disproportionately heavy use, specified the proportion of overall use a minority of books typically receives, clarified the relationship between patterns of in-library use and external circulation, and studied the stability of use patterns over time, especially the frequency with which once used books continue to be used.¹

While the total body of use studies has greatly increased our understanding, it has not always been easy to synthesize disparate studies' conclusions into generalizations applicable to most large collections. In particular, three limitations exist. First, differences in units of analysis have often meant that generalizations would require disconcerting leaps of faith—often we want to know how book use behaves according to subject and have to settle for

knowing how different kinds of use resemble one another at the level of the physical piece. Secondly, we must consider the differences between in-library use and external use while simultaneously tolerating subtle differences in the ways various studies measure each of these. A final problem is whether the results of any study would be stable over time. We know that use persists at the title level, but few studies have addressed the larger issue of whether general use patterns for a collection are stable. If we knew at least that use is stable over time, we might be more apt to accept generalizations from one collection to another of from one type of use to another.

The current study, by bringing together a large number of use measures taken in one institution, attempts to fill some of these gaps in the literature and, in the process, to specify when generalizations may or may not be made with confidence. While use is measured mostly within a

Paul Metz is a Reference Librarian and Bibliographer, and Charles A. Litchfield is Retrospective Conversion Coordinator at the University Libraries, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061.

The authors gratefully acknowledge the assistance of Lisa Brown and William Dougherty of the University Libraries in gathering the data; the helpful advice of Jesse Arnold of the Virginia Tech Department of Statistics; and the thoughtful comments of Robert Broadus of the School of Library Science of the University of North Carolina.

five-month period, the study addresses the stability issue by comparing two time periods five years apart. It brings together 1987 data on holdings, circulation, in-house use of bound materials, and use of current periodicals, along with baseline data on circulation drawn in 1982 to focus on the following specific questions:

1. How do various kinds of use differ at the subject level? Specifically, how similar are in-library use and circulation patterns? How different is the use of current periodicals from other use?

2. How stable are circulation patterns, by subject, across a time period as long as five years?

3. How large a sample is the minimum size required to yield reliable estimates of use?

4. To what extent do differences in library holdings across subjects artificially affect the correlations among use measures differing in kinds of use measured, technique of measurement, or time period? (As this question emerges from the other analyses, its significance should become more apparent to the reader.)

METHODOLOGY

Virginia Tech (or Virginia Polytechnic Institute and State University) is a comprehensive land-grant university with an enrollment of 23,000. While the university offers degrees in all fields, its programs emphasize the applied disciplines of engineering, agriculture, and business. With 1.7 million volumes, the library holds membership in the Association of Research Libraries (ARL) and ranks between the middle and the third quartile on most ARL measures of size and activity. The library system is highly centralized. The only branch facilities are in art and architecture, geology, veterinary medicine, and a remote facility to support graduate programs in northern Virginia. The library's catalog, holdings records, and circulation traffic are maintained on the on-line Virginia Tech Library System (VTLS).

In 1983, Paul Metz reported the results of a large-scale study of the use of the Virginia Tech collections. While the study reported overall use by subject categories, its emphasis was on "who uses what"

and the implications of the borrower-book relationship both for library practice and for the sociology of science.²

The main data for the 1983 study came from a "snapshot" of the 58,457 books charged to 10,126 active borrowers during May 24-25, 1982. This snapshot did not record circulation transactions but identified, by subject and borrower, the books in circulation on the evening that the data-gathering program was run. Eighty-one subject categories, defined by call-number range, were used in the study.

In 1986, a management statistics package was added to the VTLS software. Among its features was the ability to record circulation transactions across thirty call-number ranges. All checkouts and item renewals at all library sites are recorded. Separate but similar statistics are kept for batch renewals, a feature that allows patrons to renew all their eligible books in a single transaction.

The ready availability of these data naturally raised interest in how circulation patterns might have changed in the five years since the original study. The present authors were allowed to specify the call-number ranges to be used in the library's first application of the management statistics package and made a point of defining their thirty categories consistent with the eighty-one used in 1982. Because the outer boundaries of the thirty call-number ranges chosen correspond to the boundaries of various ranges from the prior study, observations in the eighty-one categories of the older study could be collapsed into observations across the thirty new categories. Table 1 shows the call-number ranges selected for use in the current study.

Both the selection of subject ranges and some problems with the actual gathering of the data require explanation. The subject categories obviously vary widely in their size and significance, but the reasons for this are not so arbitrary as may first appear. In some cases it was necessary to delimit small subject categories of relatively little interest so that other categories could be defined. For example, the study team had limited interest in the use of materials in astronomy (QB), but had to define this

TABLE 1
SUBJECT CATEGORIES

Class	Subject
Dewey, A	Unreclassified Dewey materials, general
B	Philosophy, psychology, religion
C-D	General and world history
E-F	History of the Americas
G	Geography, anthropology, recreation
H-HA	General social sciences, social statistics
HB-HJ	Economics and business
HM-HX	Sociology
J	Political science
K	Law
L	Education and Virginia Tech publications, including theses and dissertations
M	Music
P-PQ	Miscellaneous language and literature
PR	English literature
PS	American literature
PT-PZ	Miscellaneous language and literature, juvenile
Q	General science
QB	Astronomy
QC	Physics
QD	Chemistry
QE	Geology
QH-QR	Life sciences
R	Medicine
S-SD	Agriculture
SF-SK	Animal science
T	Technology, engineering, home economics
U-V	Military, naval science
Z	Bibliography, library science

range in order to delimit both mathematics and computer science (QA) and physics (QC).

With a limit of thirty subject categories, not all parts of the collection could be measured in sufficient detail to specify disciplines with the desirable precision. The decision was made to sacrifice detail within class T by devoting only one category to this large and important area. Since it had required eleven subject categories defined by sixteen call-number ranges to specify the various fields of engineering plus home economics in the 1982 study, to preserve detail within class T for the present study would have required an enormous sacrifice of precision in other call-number areas.

One consequence of the reasoning that guided the selection of call-number ranges was that the ranges chosen varied enormously in their size. At the extreme, thirty-three times as much of the collection is in subject category T as is held in QB. Later portions of the analysis will ex-

plore the methodological consequences of these disparities.

Because the VTLS statistical package is somewhat less sophisticated than the programs that gathered data for the 1982 study, some minor problems were encountered. Unlike the 1982 program, the statistical package could recognize neither government documents nor unassociated records for first-time circulations. Consequently each use of a government document (which accounted for just under 1 percent of use in the 1982 study) incremented the call-number category beginning with the same letter as the SuDocs number of the book being circulated. This made it impossible to measure use of documents and caused a small degree of error throughout the data.

VTLS substitutes the character string "not yet entered" for the call number of freshly circulated items not yet associated with bibliographic records. Unfortunately the statistical package, with infamous computer literalness, incremented the use

counter in class N for each circulation transaction of these materials. Because "not yet entered's" could not be disaggregated from use of actual N materials, class N was thrown out of the study. Percentage totals for all measures reported in this study, whether of circulation, other use, or holdings total to 100 percent because of the complete elimination of class N from the data set.

Circulation measures were gathered separately for every month from January through May 1987. The May data actually report measures from May 1-16 and May 20-29 only, as the statistical package was inoperative for five days. For April and May, but not the other months, it was possible to record data for batch renewals.

The various automated measures of circulation activity were supplemented by a variety of manually recorded measures. For May 1987, shelvers in all library locations noted the call number of each of the 95,086 bound items shelved. Government documents were the only exceptions in the gathering of these data.

"A total of 26,881 observations was made of current periodicals use by subject."

During May the staff also tallied the use of the 9,468 current periodical titles on display at all library locations. Branch librarians reported that patrons in their less formal settings reshelve most of their own current periodicals and that use of current periodicals in the branch literatures of art, architecture, geology, and veterinary medicine would therefore be underreported.

The last measures taken to form the study data were holdings measurements. Linear inches of the shelflist were measured as an index of the library's holdings in each call-number range.

The final variable, in-library use of bound materials, represents in-house use of all materials other than current periodicals,

materials in class N , and documents. It includes all use of bound periodicals, which are not allowed to circulate at Virginia Tech. The variable was created by subtracting the number of April checkouts in each subject range (which totaled 32,408) from the number of bound items reshelved in May. Without this or a similar corrective measure, the reshelving statistics would have reflected a large percentage of returns from circulation and would not accurately represent in-library use. As will be shown later, one month is a good surrogate for another in the selection of circulation data. April was chosen because it is the month closest to May, for which data were not complete, and because many books charged out in April would have been returned and then reshelved in May. This means of transforming the data seemed the most logical way to derive an imperfect, but arguably very close, measurement in-library use of bound materials for May.

Because of the problems encountered in measuring the use of government documents, branch periodicals, and materials in class N , it is not possible to describe the distribution of kinds of use with perfect accuracy. However, the data in table 2 should quite closely represent the overall mixture of materials use for May 1987. They indicate that at Virginia Tech, circulation accounts for less than one third of overall use, while current periodicals account for more than a fifth of use. As Robert Broadus has noted, various studies have yielded highly disparate ratios of in-library use to circulation, depending on measurement techniques and details of the local setting.³ The present data suggest a ratio of more than two to one between in-library use and circulation if browsing of current periodicals is included and less than two to one if such use is excluded.

FINDINGS

In-Library Use Versus Circulation

Various studies of the relationship between in-library use and circulating use differ in the degree of similarity that is found between these two kinds of use. This debate is not extreme, in that no one argues that in-house use and circulation

TABLE 2
DISTRIBUTION OF KINDS OF
LIBRARY USE, MAY 1987

Type of Use	%	n	Notes
In-library use of bound materials	49.8	62,678	May 1987 shelving minus April 1987 on-line circulation transactions
In-library use of current periodicals	21.4	26,881	
Circulation	28.8	36,175	May 1987 circulation transactions × 31/26 to correct for missing data

Excludes May 1987 reshelving of 1,458 government documents separately reported.

are totally dissimilar, and no one claims that they are identical. Only the degree of similarity is at issue. Differences in the two kinds of use may be attributed either to differences in the information content of the materials themselves—novels will be checked out, amortization tables will not—or to differences among disciplines in the ways in which materials are used. Common wisdom holds that scientists rely disproportionately on a more recent literature that appears mainly in periodicals, especially recent periodicals, while the humanities place greater emphasis on retrospective materials in monographic form, though citation and use studies have yielded surprisingly little evidence on this point.⁴ One would therefore expect to find a higher ratio of in-library use to circulation in the sciences and technology than in the humanities.

The present data make it possible to calculate correlations helpful in specifying the relationship between in-library use and other use. Before entering that analysis, it seems appropriate to present the raw data and discuss the qualitative differences in use patterns among disciplines that set an upper limit on the similarity between circulation and various kinds of internal use. Table 3 gives the percentage of use attributable to each call number for current periodicals use, all reshelving, in-house use (reshelving minus April check-outs), and circulation (including batch renewals) for May 1987.

A number of pronounced differences in the kinds of materials preferred in the various disciplines are readily apparent. While these are generally in the direction that one might predict, the data help to specify these differences with some accu-

racy. One obvious example is the extent of use in call-number ranges "Dewey" and A and Z. Although the library has modest holdings of books not reclassified from the Dewey system but no periodicals in Dewey, both the percent of current periodicals use and of all in-house use in Dewey and A greatly exceed circulation in this range. In the case of the A's, much of the difference is explained by reader interest in current, general periodicals. For both subject categories, much of the explanation for this difference stems from the high proportion of materials in classes A and Z that are of a reference nature and may be used only within the building.

"Current periodicals in business and economics (HB=HJ) account for about half again as much overall use as do bound materials in this area, reflecting in part the importance of current awareness in business and finance."

Other differences in the use of current periodicals and bound materials are more revealing of intrinsic differences in disciplines and their literatures. Current periodicals in business and economics (HB-HJ) account for about half again as much overall use as do bound materials in this area, reflecting in part the importance of current awareness in business and finance. Current periodicals use is also disproportionately heavy in music (M), though the experience of the reference staff is that at least some of this arises from use attributable to students reading *Rolling Stone*, *Gui-*

TABLE 3
SUBJECT DISTRIBUTION OF
USE, BY KIND OF USE, MAY 1987

	A. Current	B. Bound	C. House	D. Circulating
Dewey, A	8.29%	2.26%	3.27%	.70%
B	2.47	4.70	4.27	6.16
C-D	3.20	2.82	2.06	4.66
E-F	1.18	2.80	1.93	3.41
G	2.92	2.09	1.35	2.76
H-HA	.72	1.10	1.42	.52
HB-HJ	18.56	11.58	12.71	10.48
HM-HX	3.66	5.72	5.26	6.50
J	1.18	1.56	1.72	1.45
K	1.21	2.19	2.77	.98
L	5.64	3.41	3.05	4.16
M	3.36	.97	.78	1.14
P-PQ	2.15	3.47	3.06	4.94
PR	.20	3.68	3.06	5.66
PS	.28	2.79	1.68	4.57
PT-PZ	.05	1.16	.54	1.97
Q	3.42	1.92	2.52	.61
QA	4.83	5.51	5.38	5.98
QB	.39	.24	.16	.28
QC	1.57	2.75	3.02	2.49
QD	2.75	3.20	3.95	1.83
QE	.25	.75	.44	1.20
QH-QR	3.49	4.85	5.10	4.24
R	3.38	4.84	5.29	4.21
S-SD	2.67	2.42	2.68	1.81
SF-SK	1.57	1.42	1.62	.87
T	18.08	16.80	16.94	15.25
U-V	1.28	.79	.81	.80
Z	1.27	2.21	3.18	.38
n	26,881	95,086	62,678	41,581

A. Current = Shelving of current periodicals.

B. Bound = All shelving of bound materials.

C. House = In-house use of bound materials (May shelving from column B minus April online circulation transactions).

D. Circulating = All May 1987 circulation transactions, online and batch.

tar Player, and similar publications.

The use of current periodicals in literature is at the opposite extreme from use in business. Use is very light and, in the case of call numbers from *PR* through *PZ*, virtually negligible. These data suggest that current awareness is a far less salient reader strategy in literature and literary criticism than in other fields.

Differences between the internal use of current periodicals and of bound materials are especially noteworthy when we consider the large proportion of in-library use attributable to current periodicals usage. The current data suggest that with 26,881 recorded uses of current periodicals and 62,678 other kinds of in-library use, current periodicals account for about 30 percent of overall in-library use.

Comparisons among columns *B*, *C*, and *D* support and help to specify other common generalizations about literature use. Since Column *B* reflects both in-library use and circulation, it is hardly surprising that with one minor exception (*U-V*), the percentage in *B* are intermediate between those for column *C* (in-library use) and column *D* (circulation.)

In his 1983 study, Metz found that humanities materials accounted for a greater proportion of circulation than of in-house use. In-house use was somewhat disproportionately high in science and technology, while in-house use and circulation in the social sciences were roughly equivalent.⁵ This finding suggests the greater dependence of scientists on journals, which in the case of the Virginia Tech libraries do

not circulate. Comparisons between the data on in-library use and circulation use in table 3 support the same conclusion Metz reached in 1983. For classes *B-F* and *M-P*, circulation percentages exceed percentages of in-house use by a considerable margin (though raw circulation is less than in-house use for virtually all categories). For most fields of science and technology, the in-house-circulation relationship is reversed in that in-library use percentages are higher than percentages for circulation. Overall, materials in classes *Q* through *T* accounted for 47.1 percent of in-house use of bound materials and 42.4 percent of current periodicals use, but only 38.8 percent of overall circulation.

If there is any surprise in this comparison, it may be that in-house use does not exceed circulation more dramatically than it does in science and technology, and especially that current periodicals use is not higher. However, both Metz and Harry Kriz have found that, common assumptions notwithstanding, scientists and engineers make substantial use of circulating materials.⁶

While in-library use of bound materials and use through circulation differ in ways that reveal something about humanities, science, and other disciplines, table 3 shows that as a rule subjects tend to vary together in their tendency to receive lighter or heavier use, regardless of which kind of use is being considered. Generally use studies comparing in-house use with circulation have found a positive correlation. An early example is Herman Fussler and Julian Simon's conclusion that "Books that develop little recorded use develop little browsing, and books that develop much recorded use develop much browsing, except for the highest use books, for which extrinsic factors distort the picture."⁷ Anthony Hindle and Michael Buckland's work supports this conclusion. Their comparison of in-house use and circulation at the University of Lancaster library led them to state that "the conclusion is very clear. Books that circulate little get relatively little in-house use and the higher the circulation the higher the level of in-library use."⁸

The most recent validation of the similarity of in-house use and circulation comes from the University of Pittsburgh study. Comparing data on the internal use of books gathered in 1975-76 and their main database reflecting circulation activity, Stephen Bulick and others found that at least 75 percent of books used in-house had also circulated externally, and that books used in-house had a higher mean number of external circulations. Their findings led them to embrace Fussler and Simon's earlier conclusion that recorded use is a reasonably good index of all use.⁹

If individual books tend to be either heavily or lightly used both internally and through circulation, the overall correlations comparing these two kinds of use within subject categories will be positive. Because the data have already shown that in-library and circulation use are by no means proportionate across subjects, the correlations, at least in the present case, must be well short of unity. How strong are the correlations, and what do they tell us about the similarity of in-library use and circulation across subject categories? William McGrath's 1971 study comparing the two kinds of use at the University of Southwestern Louisiana yielded correlations across his subject categories of .84 and .86, depending on small methodological variations. He concluded from this that "circulation totals, when grouped into self-delineating spans, can be reliable indicators of the subjects being used within as well as out of the library."¹⁰

The present data support this conclusion and, indeed, provide higher estimates of the relationship between in-library use and circulation. By measuring both kinds of use in slightly different ways, they also help to identify the degree of difference among various kinds of use. Table 4 reports the Pearson Product Moment Correlations among a variety of measures of external and internal use.

The single best estimate of the relationship between in-library use and circulation is the correlation of .885 between the variables measuring all circulation transactions for May and the May reshelving of bound materials not returning from circulation. This correlation, which is consis-

TABLE 4
CORRELATION COEFFICIENTS
AMONG MEASURES OF IN-LIBRARY
USE AND CIRCULATION, MAY 1987

Online						
Batch	.942					
Allcirc	.995*	.970*				
House	.860	.911	.885			
Current	.733	.789	.758	.902		
Shelving	.897*	.933*	.918*	.988*	.932*	
	Online	Batch	Allcirc	House	Current	Shelving

Online = Circulation online checkouts and renewals

Batch = Circulation batch renewals

Allcirc = Sum of online and batch circulation

House = May shelving of bound materials used in-house (May bound shelving minus April online circulation transactions)

Current = Current periodicals shelving

Shelving = All in-house shelving, including returns from circulation and current periodicals.

*Denotes correlation of two measures, one of which is a subset of the other.

Note: All correlations reported in this study are statistically significant at $p < .001$.

tent with McGrath's observed correlations of .84 and .86, would indicate that about 78 percent of the variance in each variable would be synchronous with variation in the other. The differences in the various kinds of use observed in table 3 make it difficult to describe in-library use and circulation as identical. Still, the two kinds of use are very similar, and, after allowing for fairly predictable differences, one may generally use one kind of use as a surrogate for the other.

"Not surprisingly, the correlations between circulation and the internal use of current periodicals are lowest of all those reported here."

The data in table 4 show important differences in the kinds of in-library use recorded in this study. The correlation of .902 between the subject distribution of in-library use for bound materials and current periodicals is high, but does reveal a sufficient difference to indicate that the distinction between these kinds of internal use deserves a degree of attention not hitherto received in use studies. Not surprisingly, the correlations between circulation and the internal use of current periodicals are lowest of all those reported here. Future statements about the similar-

ity of circulation and in-house use should carefully specify whether the use of current periodicals has been counted.

While the distinctions among kinds of internal use are important, those between circulation measures are not. Differences in the subject distribution of books being checked out or renewed at the item level and those being batch renewed are relatively unimportant ($r = .942$).

Stability of Circulation Patterns

An unusual feature of the present data is the time period covered. Few if any studies have been able to compare use data drawn at points as far apart as five years. These data make it possible to investigate the stability of circulation measures, first across measurement periods close together and then across a time interval of a half-decade.

Table 5 reports the Pearson correlation coefficients for measures of circulation transactions across the twenty-nine subject categories for the individual months January through May 1987 and for the sum of these periods. The lowest correlation between pairs of individual months is .967. The lowest correlation of any month with the total period is .987. These data require little interpretation. They clearly show that circulation data by subject for any month are essentially interchangeable with those for any other and would provide an excellent measure of overall circu-

TABLE 5
CORRELATION COEFFICIENTS, JANUARY
THROUGH MAY ONLINE CIRCULATION TRANSACTIONS

JANTRAN						
FEBTRAN	.973					
MARTRAN	.977	.993				
APRTRAN	.987	.989	.990			
MAYTRAN	.967	.994	.990	.984		
ALLTRAN	.987	.996	.996	.996	.993	
	JANTRAN	FEBTRAN	MARTRAN	APRTRAN	MAYTRAN	ALLTRAN

TABLE 6
CORRELATION COEFFICIENTS,
SELECTED 1982 AND 1987 CIRCULATION MEASURES

ALLTRAN87				
APMAYALL87	(.996)			
MAY82SNAP	.968	.976		
OCT82SNAP	.958	.971	(.991)	
	ALLTRAN87	APMAYALL87	MAY82SNAP	OCT82SNAP

- ALLTRAN87 = All January-May 1987 circulation online transactions
 APMAYALL87 = All April and May 1987 circulation transactions, online and batch
 MAY82SNAP = May 1982 snapshot of books in use
 OCT82SNAP = October 1982 snapshot of books in use

lation. Provided that school is in session and use of a library is normal, the issue of variation in use patterns by subject through the academic year appears to be irrelevant.

The data in table 6 represent the correlations among two pairs of circulation measures taken five years apart. The variable "ALLTRAN87" represents all online circulation transactions for January through May 1987. "APMAYALL87" represents the sum of both online and batch circulation transactions in April and May 1987. "MAY82SNAP" and "OCT82SNAP" represent the main 1982 data, a snapshot of books charged out in May and an October 1982 snapshot taken to assess the stability of the circulation data between spring and fall 1982.

In the table the correlations between measures taken in the same year are reported in parentheses so as to highlight the four correlation coefficients reporting use across the five-year interval. All four correlations are very high, despite both the five-year interval and differences in measurement technique among online transactions, the sum of online transactions and batch renewals, and snapshots of materials in use.

These findings buttress the argument that circulation measures are relatively immune to differences in measurement technique. More importantly, they reveal an extraordinary measure of stability, indicating that circulation patterns in an institution not undergoing dramatic curricular change or extensive changes in the direction of library acquisitions are remarkably constant. Together with the data in table 5, these correlations suggest that a single measure of circulation use drawn at any normal time in a given school year may provide an academic library with measures on which it can rely for a considerable period of time.

Use of Small Samples to Assess Collection Use

The knowledge that circulation patterns are as stable as the current data suggest may encourage more academic libraries to conduct circulation studies. While such studies are relatively simple to conduct in automated libraries, they may be laborious and expensive in libraries whose circulation procedures are manual. This raises the question of cost-benefit trade-offs in the selection of sampling periods, and specifically the question, What is the

shortest (hence cheapest) period that will yield reliable estimates of circulation activity?

The team of researchers that studied library use at the University of Pittsburgh in 1975 anticipated this question and conducted their own research toward answering it. By drawing *ex post facto* subsamples of their data and comparing these to the full data set, they were led to conclude that "random samples of as few as three days produced correlations as high as .95-.98 with the total population in regard to percent of usage by LC class."¹¹

To replicate the work of the Pittsburgh team, data were drawn from the present data set to provide subsamples of time periods equal to and both longer and shorter than the three-day period found sufficient by the Pittsburgh team. The analysis has already shown that one-month samples more than adequately represent circulation activity for a longer time period. The subsamples drawn to refine this conclusion further consisted of two seven-day periods from April 1987; two three-day periods from April; and two single days drawn from March. The correlations of these subsamples with the entire set of 160,141 online circulation transactions are shown in table 7.

The results both replicate the Pittsburgh findings and indicate that the three-day period they found adequate may well define the lower limit for short sample periods. The results for three-day subsamples from the full data were as strong as those found at Pittsburgh. Indeed, three-day samples yielded results as reliable as those from seven-day samples. However,

correlations noticeably dropped off when subsamples of only one day were used.

In interpreting these findings, it should be remembered that the issue of sample size is at least as important as that of sampling time periods. A small library with modest levels of circulation would risk obtaining unreliable estimates from a three-day sample because of insufficient volume of circulation transactions and the possibility that one or two individual patrons with many books could unduly influence the entire data set.

Use Measures and the Effects of Library Holdings

To some degree the use of any library collection must reflect the self-fulfillment of prophecies made by those who built the collection. Stated negatively, patrons cannot use materials the library never obtained. Stated positively, patrons may find their reading interests influenced by the strengths of a given library collection.

If the profile of a library collection does influence use, it must necessarily influence all kinds of use over various time periods. This raises the question of whether the correlations among various measures differing in kind of use, measurement technique, or time period may not be artificially boosted by the artifact of an individual collection's profile. A more detailed discussion of this question is given in the appendix.

While this question is troubling, it is not intuitively obvious that it is a serious one, or that one is not "borrowing trouble" even to raise it. However, a conscientious approach to the current analysis requires

TABLE 7
CORRELATIONS OF SELECTED
SUBSAMPLES OF ONLINE CIRCULATION TRANSACTIONS
WITH ALL ONLINE CIRCULATION TRANSACTIONS

Subsample	Correlation Coefficient
April-7-A	.989
April-7-B	.987
April-3-A	.989
April-3-B	.991
March-1-A	.922
March-1-B	.925

Variable names indicate month, number of days, and first (A) or second (B) randomly selected period.

that the possibility of artifactual effects on the findings be explored.

Three considerations in particular mandate the pursuit of this question. One is the independent findings by Thomas John Pierce and William E. McGrath that holdings statistically explain about 65 percent of the variation in use across subjects.¹² That is, correlations between holdings and use across subjects have been found to be about .8. The second consideration is the large disparities in size between holdings across the subject categories used in the current study—as mentioned, the subject categories vary by a factor of thirty-three in the proportion of the collection for which they account.

The final consideration requiring some investigation of this issue is the realization that the correlations reported in the study are consistently somewhat higher than those found elsewhere. As one example, the correlation of .991 between the May and October 1982 circulation snapshots exceeds the correlation of .982 Metz found for the same data when he originally compared them using his eighty-one subject categories, which did not differ so dramatically in size as those used here.¹³

The relationship between use and holdings for the current study was found to be in the same range as was found by Pierce and McGrath, indicating that perhaps the larger size disparities among subjects in the current study are not an important

consideration. The correlation of all April and May online and batch circulation transactions with holdings across the subjects was .784. The correlation of all online circulation transactions from January through May with holdings was .793. Both measures are essentially identical to the Pierce and McGrath findings.

To assess the possible effects of holdings in artificially boosting the correlations among use measures, partial correlations were calculated among the use measures with statistical controls for holdings. The derivation of partial correlations represents a standard approach to identifying the relationship between two variables after accounting for the effects on each of a third variable.

The results indicated that allowing for holdings effects does considerably reduce correlations between in-library use and circulation, but that correlations among other measures of use or among similar measures taken at various time periods remain extremely high despite these controls.

Table 8 shows a representative selection of some of the more important correlations among use measures reported in this study. The drop-off in correlations between in-house use and external circulation (see first two items in table 8) is substantial, although the partial correlations of .730 and .664 remain quite healthy and explain about half the variance. Measures

TABLE 8
CORRELATIONS, ZERO-ORDER AND WITH
CONTROLS, AMONG SELECTED USE MEASURES

Measure 1	Measure 2	Raw Corr.	Partial r Controlling for Holdings
MAYALL	MAYHOUSE	.885	.730
MAYTRAN	MAYHOUSE	.860	.664
JANTRAN	ALLTRAN	.987	.975
MAYTRAN	ALLTRAN	.993	.984
APRIL-3-A	ALLTRAN	.989	.970
APRIL-3-B	ALLTRAN	.991	.976
MAY82SNAP	APMAYALL87	.976	.937

MAYALL

MAYHOUSE

JANTRAN, MAYTRAN, ALLTRAN

APRIL-3-A/B

MAY82SNAP

APMAYALL87

= All May 1987 circulation transactions, both online and batch

= All May 1987 in-house use of bound materials

= Online circulation transactions for January, May, and January-May

= Three-day subsamples of April circulation transactions

= May 1982 snapshot of books in use

= All April and May 1987 circulation transactions, online and batch

of the stability of circulation across time and across different techniques of measurement are much more robust and show less diminution from these statistical controls.

For those correlations that are significantly reduced by controls for size of holdings, we should remember that such controls have not previously been introduced in the literature. The raw rank-order correlations remain the best single measure of the consonance between indicators of use and should be compared to similar correlations elsewhere in the literature. However, their reductions with statistical controls for holdings do indicate that the effect of holdings on use is a valid concern and should raise the issue of what reduction such controls would bring to the correlations reported by other studies. To the extent that the subject categories used in other studies may not vary in size so widely as those used here, there will be less need for concern about the spurious effects of holdings. Still, it would be prudent for others conducting research in this area to assess the effects of holdings within their own studies.

For those correlations relatively immune to such statistical controls, our confidence should be increased in these findings about the stability of use measures over time and across various measurement techniques.

CONCLUSION

While the current study has investigated a broad range of issues, all the questions it raises treat the comparability of use measures across one of three dimensions: types of use, measurement techniques for assessing use, and time. We have essentially asked, How stable and how similar are measures of use? in a variety of ways. The following statements summarize our several responses to this question.

1. In-library use of bound materials and

circulation are highly correlated. The differences between them follow fairly predictably from differences in the kinds of information found in various call-number ranges and from the disproportionate use of circulating materials in the humanities as opposed to the sciences.

2. In-house use of current periodicals is qualitatively different from other kinds of use, though it does correlate significantly with them. Future studies should devote more attention to this kind of use, which accounts for a large fraction of in-library use, and should carefully distinguish it when results are reported.

3. Circulation statistics are extraordinarily stable over time and extraordinarily insensitive to minor differences in measurement techniques.

4. For libraries with large circulation volume, three-day samples of circulation activity yield highly satisfactory estimates of overall circulation volume.

5. Holdings influence library use and introduce a factor that future studies are obliged to consider. Holdings account for a substantial part of the observed relationship between in-library use and circulation. They do not appear to have so important a role in accounting for the observed relationships among other measures of use.

It has become traditional to end studies of this nature with calls for replication. Indeed, replication of this study's novel findings, especially concerning the nature and extent of current periodicals use, is highly desirable. Still, in important respects this study replicates others. By showing the stability of circulation data, it also suggests that most use studies will be readily replicable within their own settings. Taken together, these findings may caution against a paralyzing overemphasis on the importance of replication in studies of library use.

REFERENCES AND NOTES

1. Allen Kent and others, *Use of Library Materials; The University of Pittsburgh Study* (New York: Marcel Dekker, 1979); Herman Fussler and Julian Simon, *Patterns in the Use of Books in Large Research Li-*

- braries (Chicago: Univ. of Chicago Pr., 1969); Michael K. Buckland, *Book Availability and the Library User* (New York: Pergamon, 1975); Richard W. Trueswell, "User Circulation Satisfaction vs. Size of Holdings at Three Academic Libraries," *College & Research Libraries* 30:204-13 (May 1969); William E. McGrath, "Correlating the Subjects of Books Taken Out of and Books Used within an Open-Stack Library," *College & Research Libraries* 32:280-85 (July 1971).
2. Paul Metz, *The Landscape of Literatures: Use of Subject Collections in a University Library* (Chicago: American Library Assn., 1983).
 3. Robert Broadus, "Use Studies of Library Collections," *Library Resources & Technical Services* 24:317-24 (Fall 1980).
 4. Citation studies showing the greater dependence of the natural sciences than of other disciplines on recent publications include Derek J. deSolla Price, "Citation Measures of Hard Science, Soft Science, Technology, and Nonsense," in ed. Carnot E. Nelson and Donald K. Pollock, *Communications among Scientists and Engineers* (Lexington, Mass.: D. C. Heath, 1970), p.3-22; and Duncan Macrae, Jr., "Growth and Decay Curves in Scientific Citations," *American Sociological Review* 34:631-35 (Oct. 1969).
 5. Metz, *Landscape of Literatures*, p.23.
 6. Metz, *Landscape of Literatures*, p.14; Harry M. Kriz, "Subscriptions vs. Books in a Constant Dollar Budget," *College & Research Libraries* 39:105-109 (Mar. 1978).
 7. Fussler and Simon, *Patterns in the Use of Books*, p.115.
 8. Anthony Hindle and Michael K. Buckland, "In-Library Books Usage in Relation to Circulation," *Collection Management* 2:265-77 (Winter 1978).
 9. Stephen Bulick, William N. Sabor, and Roger R. Flynn, "Circulation and In-House Use of Books," in Kent and others, *Use of Library Materials*, p.9-55.
 10. McGrath, "Correlating the Subjects," p.285.
 11. Bulick and others, "Circulation and In-House Use of Books," p.44.
 12. Thomas John Pierce, "The Economics of Library Acquisitions: A Book Budget Allocation Model for University Libraries," (Ph.D. diss., Univ. of Notre Dame, 1976), p.46; William E. McGrath, "Predicting Book Circulation by Subject in a University Library," *Collection Management* 1:7-23 (Fall/Winter 1976-77), p.12.
 13. Metz, *Landscape of Literatures*, p.116.

APPENDIX A. The Possible Effects of Holdings on Use

This issue may be best clarified in the form of a debate.

PRO: Collections will influence use in a way that artificially heightens correlations among use measures.

If monkeys were turned loose on the Virginia Tech collection on Monday and gibbons roamed the stacks on Tuesday, both would disturb many more books in class T than in class QB simply because there are so many more of the former. This would yield high but meaningless correlations between Monday and Tuesday's use. Only those books actually in the library can be used, and therefore use is to some degree an outcome of holdings.

CON: The influence of collections on use patterns does not introduce a significant methodological artifact.

Holdings may "explain" about 65 percent of use in the narrow statistical sense, but do they substantively "explain" anything? Some subjects attract more interest than others. More books are published in them, libraries buy more books in them, patrons read more books in them. If the correlations between holdings and use were zero, library collection development activities would be condemned as random processes. Statistical measures aside, the issue of just what is explained by holdings is an open one.

We should remember that we are not considering the use patterns of monkeys and gibbons, but of students and faculty, who will select only those materials they find useful. The existence of materials in a collection does not make their use inevitable. If for some reason 10 percent of the Virginia Tech collection were in numismatics, would one expect that these materials would receive 10 percent of use?