Suggestions for Statistical Records, II

A discussion begun in the June 1945 number of College & Research Libraries is concluded in the following pages.

By FAR the weakest category in the Library Statistical Report for Institutions of Higher Education is the rubric headed "Number of Photostats or Microfilms," which is broken down into the number made for other libraries and the number obtained from other libraries. This figure is as useless as though the Boston fish market were to lump together the daily totals of lobsters and mackerel.

We need to start over again in keeping records of this extremely important branch of modern library science. While it was still relatively undeveloped when the Library Statistical Report for Institutions of Higher Education was taking shape, it is nevertheless a service which might well become the most important single library activity within the foreseeable future; and, accordingly, it deserves better treatment than it has been receiving from most libraries. In the first place, we need precise information as to the nature of the equipment in each library and what personnel is available to operate it. A description of the equipment might fit into the suggested quinquennial issues of the statistical handbook previously mentioned. Annual records of work done should be noted under the following headings: number of microfilm exposures, number of microfilm prints, number of photostatic exposures, number of photographic negatives, number of contact prints, and number of enlargements from microfilm. A very large laboratory, such as that maintained by the Library of Congress' Photoduplication Service, needs additional figures on such items as color transparencies, lantern slides, projection prints, blueprints, Ozalid prints (square feet), multilith plates, and multilith (copies). Figures on materials received from other libraries are needed for microfilm prints, photostatic exposures, and contact prints. Microfilm can be expressed either in feet or in numbers of exposures.

It should be noted that photographic services are intimately related to interlibrary loan. To a large extent, we in America are supplanting certain types of interlibrary loan with photostats and microfilms. While the European librarians assiduously collect statistical records of international interlibrary loan and of interlibrary loan of manuscripts, we refrain from gathering this information simply because the services themselves are negligible, due to our policy of discouraging such loans in favor of photographic work and because our holdings of papyri,


26. Marcel Godet's complaint in his "Le Prêt International des Livres et Manuscrits" (Der Schweizer Sammler 85-96, 113-23, 1937) that international interlibrary loan statistics are neglected, can be answered by American libraries merely by listing in separate categories figures for photographic work done for libraries and individuals abroad.

27. Muszkowski, J. "Interlibrary Loans in Poland." Publications of the International Federation of Library Associations 5, 1934—Actes du Comité International des Bibliothèques, 6ème session, Chicago, October 14, Avignon, Nov. 13-14, 1933, p. 196, argues that photography will not take the place of international interlibrary loan inasmuch as "it cannot be recommended to reproduce all pages of a thick volume, if the borrower is especially interested in one or two pages only." A careful scholar will usually have an exact page or leaf reference or can give directions for ascertaining it readily.
ancient, medieval, and renascence manuscripts are, relatively, far smaller.

Interlibrary loan figures called for on the Library Statistical Report for Institutions of Higher Education tell us only how many volumes are borrowed and how many are lent. Possibly because these figures mean so little they were omitted from the published tables in College and Research Libraries. Certainly their chief virtue is a negative one—they are expressed in terms of volumes rather than in terms of “packages” or “shipments,” which some libraries have used. Still, the simple figures of books borrowed and lent have some value as a means of ascertaining trends and determining the true growth of a library as reflected in its ability to supply its own readers adequately without having to request loans from the outside. For example, twenty years ago the Iowa State College Library borrowed far more books than it lent; but, after two decades of carefully supervised book selection, its “balance of trade” in interlibrary loan is almost exactly reversed.

In the first place, it would seem desirable to treat interlibrary loan statistics just as any other circulation statistics (infra) and to analyze the books lent by classes. In the second place, it would be useful for research libraries to divide institutions from which books are borrowed into (a) Library of Congress, (b) other college and university libraries, and (c) other libraries; and institutions to which books are lent into (a) other college and university libraries, (b) public libraries, and (c) school libraries. These figures would be of considerable value as a guide to qualitative evaluation of the significance of interlibrary loan in a particular institution. There can hardly be any serious objection to compiling these figures when many institutions consider it worth their while actually to list all institutions to which they lend books.

Another type of figure on interlibrary loan is represented in a table constructed by Constance M. Winchell giving “Statistics Showing Location Found for Books Requested for Interlibrary Loan at Columbia”:

<table>
<thead>
<tr>
<th>Year</th>
<th>1927-28</th>
<th>1928-29</th>
<th>1929-30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of requests</td>
<td>305</td>
<td>378</td>
<td>482</td>
</tr>
<tr>
<td>Located through records at Columbia</td>
<td>167</td>
<td>223</td>
<td>316</td>
</tr>
<tr>
<td>Located by letter</td>
<td>103</td>
<td>73</td>
<td>122</td>
</tr>
<tr>
<td>Not located</td>
<td>35</td>
<td>82</td>
<td>44</td>
</tr>
<tr>
<td>Total located</td>
<td>270</td>
<td>296</td>
<td>438</td>
</tr>
</tbody>
</table>

In passing, it might be noted that many a reference librarian has dreamed of his secret

---

weapon as statistical records of bibliographically incomplete requests.

Perhaps the most difficult of all statistical records to collect are those on reference work. From almost any angle that one approaches the problem, it seems so. The Library Statistical Report for Institutions of Higher Education asks for the number of general questions answered, the number of search questions answered, and the number of bibliographies compiled. The “Daily Statistical Record” and “Monthly Statistical Record” used by the library of Teachers College, Columbia University, classifies reference questions into “location,” “easy,” and “requiring time.” While many defects may be found in the categories used by both of these forms, no better solution can be offered here at present. Nevertheless, this matter deserves careful study and would seem to be well worth a master’s essay from one of our library schools.

Counting attendance in the main reading room or in the entire library is rather futile except in such institutions as the J. Pierpont Morgan or Henry E. Huntington libraries, where all readers are known to be in the library for a serious purpose. The New York Public Library Reference Department follows an interesting procedure for counting readers in the main reading room. The number of books used by means of the submission of signed call slips is divided by an arbitrary figure, varying from 2.07 to 2.38, depending on the time of the year. This figure is based on figures accumulated over a period of many years, at a time when the number of readers in the main reading room was recorded by a check of the signatures on the call slips day by day.

This procedure is not without significance for university libraries, especially those which have large reading rooms confined to special disciplines and in which circulation is confined to the building. But if indiscriminate attempts are made to count total attendance, such a large number of factors—weather, examination periods, puzzle contests, and even fraternity initiations—must be considered, that the results will be hopelessly confused.

Before approaching the problem of circulation statistics, it might be pertinent to repeat a delightful story, told by Henri Lemaître, about the methods of an elderly French librarian who was instructed to keep circulation statistics. “It’s very simple,” he explained. “The first time they asked me for this figure, I calculated it something like this: every day I issue about ten volumes, and since I am open 250 days a year, the result is 2500 volumes circulated. In order to appear a spot more exact, I put the figure at 2467. Since then I have increased it discreetly every year.”

Value of Present Methods

The question before us is not whether the incorrigibly inexact old Frenchman should have adopted more scientific procedures but whether our methods presently in use are much better than his. Today we may be grateful to the Library Statistical Report for Institutions of Higher Education for having set forth several definitions and rules regarding circulation records which are distinct advances over categories used in other reports and collective compilations.

Amount of Use of the Reference Department of the New York Public Library. Bulletin of the New York Public Library, 40:909, November 1936. This entire article, consisting of p. 907-25, may be highly recommended as an example of practical use of statistical records. Recently, samples of actual attendance have been taken in the main reading room of the New York Public Library to ascertain whether these “arbitrary figures” are still valid. It will be interesting to learn whether they are still valid.

in recent times. The treatment of renewals, definition of “volume” as applied to circulation, rules on counting, and regulations for recording loans from one agency to another within the same system, are of greatest value. However, no system of reporting circulation statistics from a large number of libraries has ever yet been devised which takes into due consideration such factors as varying periods of loan, restrictions on the loan of certain classes of books, effect of hours of opening, varying types of departmental collections, use of outside libraries by students, and local conditions ad infinitum. Carl M. White has stated the whole problem of circulation statistics quite accurately: "The story that circulation statistics tell us is incomplete and, to a certain extent, unilluminating. Incomplete because confined strictly to recorded circulation; unilluminating because confined to results, with no clue as to what produced them." 34

A good example of why circulation figures are not comparable is furnished by the statistical records of two institutions which are comparable in other respects and whose libraries are known to keep dependable statistical records. We learn from the U. S. Office of Education's College and University Library Statistics, 1939-40 that, in 1939-40, Duke University reported total library holdings of 600,235 volumes, an enrolment of 3,238 undergraduates and 298 graduates, and a circulation of 72,549 volumes for home use and 144,820 reserved books. The University of North Carolina, an institution with similar aims and purposes as far as the instructional program is concerned, reported a total of 386,390 volumes in its stacks, an enrolment of 3,180 undergraduates and 569 graduates, and—perhaps, for all the uninformed might know, in revenge for a decade of humiliating football defeats by Duke—a circulation of 368,071 volumes for home use and 173,006 reserved books. Somewhere in the administrative records of these two libraries there is a clear explanation for this fantastic discrepancy in their circulation figures, but this doesn't help the student of library administration who wants to compare and tabulate these figures to solve some problem of library use.

**Comparable Figures**

Even though it may never be possible to collect absolutely comparable circulation statistics, it is possible to classify circulation by subject and thus to study the comparative circulation of books in various subject fields within one particular library. This is no new idea, but it is a good one. While William F. Poole recommended it in the *Special Report* of 1876, few libraries have followed his advice except by segregating circulation of fiction and non-fiction. The New York Public Library's splendid statistical reports are particularly valuable for their excellent analysis of circulation by the various subject fields. An especially convincing plea was made for information on library use by specific departments in colleges by E. W. McDiarmid, who pointed out its value in ascertaining the extent to which students in each department use the library and in apportioning the book fund. 36

---


35 White, Carl M. "Trends in the Use of University Libraries." *Ibid.*, p. 10. This type of study is about the only kind which can be validly based on library circulation figures as gathered in the United States. For further notes on the deceptiveness of circulation statistics see the U. S. Office of Education, *College and University Library Statistics*, 1939-40, p. 9, 104, note 5.


A particularly important problem is the question of delivery service at the circulation desk. American compilations of library statistical records shed no light on this most important subject. Some individual libraries do record this information, and the *Jahrbuch der Deutschen Bibliotheken* gives it for German libraries. In the *Jahrbuch* there is a column for the total number of call slips submitted in each library, followed by four columns giving the following disposition of call slips, together with the percentile relation of each type of disposition to the whole: (a) delivered, (b) in use, (c) not available, and (d) not in the library. The latter category is valid only for European libraries, where the public does not have access to the catalog. The third category is inexcusable. A book may be “in bindery,” “missing since —,” “in process,” or even in the “inferno”, but it should never be “not available” to the appropriate reader at the appropriate time.37

The significant points are to show (a) average and possibly also median time needed for delivery, as taken from representative samples, and (b) books not delivered, why not delivered, and time consumed before readers receive answers, also based on representative samples. The land-grant college survey used a highly suggestive form to secure this information.38 Here are data from which we can hazard some kind of comparison of library service and draw some definite conclusions as to openings for improvement.

A final delinquency in the statistical records of circulation in American libraries is that we don’t know who uses the library. The *Library Statistical Report for Institutions of Higher Education* requests in-
University Libraries, in which the readers in the Hoover Library are analyzed into the following categories: staff and research workers, faculty, graduate students, undergraduate students, and visitors. Even though it is not practical to ascertain the relative amount of student and faculty use of a library as a whole, it may be determined for special collections with separate reading rooms.

There has been considerable attention paid to the problem of undergraduate use of the library.\(^40\) An investigation of this matter should give special attention to reserve book loans, noting (a) the number of withdrawals and (b) the number of titles withdrawn. The proportion of books placed on reserve, which are actually used, is also revealing in this connection.\(^41\)

**Cataloging Statistics**

For some reason virtually all compilations of statistical records and most annual reports of libraries choose to ignore cataloging statistics.\(^42\) Somehow or another, librarians seem to be very reluctant to give out these figures. They will describe in glowing detail some handsomely printed catalog of incunabula which they have just published, but are reluctant to divulge the total number of more commonly used books which have been made accessible to the public. Perhaps they feel that the recent admirable studies on library cost account-

\(^40\) The more significant studies are noted in chapter 2 ("How Much Do Undergraduates Use the Library?") of Branscomb, Harvie. *Teaching with Books.* Chicago, A.L.A., 1940. Especially important is McDermid, E. W. "Conditions Affecting Use of the College Library," loc. cit.

\(^41\) See Hurt, Peyton. *The University Library and Undergraduate Instruction; An Analysis of Their Relationship.* Berkeley, University of California Press, 1936, p. 15.


\(^45\) Dornin, op. cit., p. 82, reports that "replacements and added copies are carefully tabulated by ten libraries, but the remaining thirty-eight count them simply as added volumes." (From a survey covering forty-eight libraries.)

\(^46\) At the Iowa State College Library, where nearly a fourth of the total holdings have been recataloged and reclassified in the last quarter of a century, it was found that recataloging and reclassifying was more expensive than processing new books. First the old cards must be pulled, and then the volume to be re-cataloged is treated like a new book.

SEPTEMBER, 1945

327
portion for which original cataloging is necessary, number of old cards to which additions are made, number of subject headings assigned, number of analytics prepared, 47 numbers of cards (or possibly numbers of inches of cards) filed in public, official, depository, and special catalogs, and arrears (if the library has no reason to hide them).

It is probably unnecessary to observe that figures of enrolment in colleges and universities are frequently heavily padded. The librarian does not make his own count of students but, rather, gets it from the registrar’s office. Any study of a university library’s statistical records in relation to student enrolment is open to question. It is hardly logical to figure out library expenditures per student by dividing the total enrolment into the library appropriation when, in an enrolment of 12,000, some 3,000 may be correspondence students and some 4,000 short course enrollees. The college or university which pads enrolment figures is only working to its own disadvantage. When such figures as per capita expenditures, number of instructors per student, number of volumes in the library per student (in some respects, but not all, a rather futile figure), are compiled in attempts to ascertain an institution’s essential excellence, it comes out far worse than it otherwise would have. 48

Conclusions

Most of the conclusions which can be drawn from this brief critique of statistical reporting in libraries of higher institutions are already apparent. Above all, it is obvious that the Library Statistical Report for Institutions of Higher Education needs revision. 49 But no amount of collecting information will be of any value unless there is adequate provision for its organization and publication in a form as complete as possible. As for the statistical categories themselves, the most immediate needs are for revision of definitions (e.g., of terms relating to acquisition and book counting), more detailed breakdowns of figures already gathered (e.g., of binding and interlibrary loan), and adoption of additional devices which can give more adequate conceptions of both quantitative (e.g., measurement of holdings in linear feet) and qualitative (e.g., analysis of holdings by subject field) aspects of a given library. 50

It might also be noted that library statistics also show certain general deficiencies in common with other government data. 51 In spite of the efforts of the A.L.A. to act as central headquarters for the collection of statistical data, there is no central source yet, since there seems to be no general policy as to where the responsibilities of the A.L.A. and the U.S. Office of Education in gathering data begin and end. Except for a few isolated cases, such as James A. McMillen’s Statistics of Southern Colleges

47 Or the time spent on analytics by the cataloger (Dornin, op. cit., p. 83).
48 Ellsworth, Ralph E. “Trends in University Expenditures for Library Resources . . . 1921-41.” Library Quarterly 14:8, January 1944, notes that university administrators are even more backward than librarians in collecting and publishing certain types of statistical records.

49 Furthermore, certain points are mentioned in this article which are pertinent for large public libraries serving research workers as well as the public at large and which might be reviewed in any revision of the “Public Library Statistical Report” (U.S. Office of Education. From 8-071, 1940) which may be undertaken.
50 Burgess, op. cit., p. 92-95, notes the advantages of using such refined statistical devices as logarithmic graphs (Hawkins, E. R. J. “Logarithmic Graphs.” Library Association Record 39:257-60, June 1937) and index numbers for expressing serial statistics with long years of continuity, such as the A.L.A. statistics. He is also very enthusiastic about the use of punched cards as a means of technical implementation (p. 95-96), a device which may be an answer to the objection of many librarians that they are already spending too much time collecting statistics. An amusing example of unnecessary work caused by ignorance of sampling techniques may be seen in “Die Benutzungweise der öffentlichen Bibliotheken,” in Körosi, Joseph, ed. Statistisches Jährbuch der Stadt Pest 1:1353-56, 1873 (also carried as Vol. 7 of Pest, Sz. K. Pestváros Statisztikai hivatalának közleményei. Publikationen des Statistischen Bureau der Stadt Pest).
51 Burgess, op. cit., p. 79-82, lists these deficiencies as based on categories set up in Gray, Edward R. “Deficiencies in State and Local Government Data.” Annals of the American Academy of Political and Social Science 207:194-204, January 1940.
The data on library statistics is scattered and not very useful. It is hard to get a clear picture of what is happening in libraries across the country. There is little consistency in the way data is collected and reported. For example, some libraries keep track of how often their catalog is used, while others might track the number of magazines on a certain table at a certain time of day. The data collected by different agencies is not always comparable, making it difficult to draw meaningful conclusions.

Despite these challenges, there are some promising signs of improvement. The A.L.A. has been more active in collecting and reporting statistics, and there are now plans to publish a statistical handbook for North American libraries. This will require cooperation from all types of libraries to represent their interests fully.

The handbook will need to be regularly updated to keep pace with changes in administrative practice. It's important to get feedback from all types of libraries to ensure the handbook is useful and relevant.

It would be desirable for each major European country to be treated in a manner similar to Burgess' analysis of sources of statistical records for American libraries and Gering's for German libraries.
Special care should be exercised not to permit the forms to become cumbersome by the inclusion of categories which are of little or no significance for the particular type of library concerned. Membership on these committees should include representatives of acquisition, circulating, cataloging, and reference departments. The present A.L.A. Committee on Statistics might be reorganized to form a higher control committee which would have less to do with the technical details of collecting and tabulating statistical records than with the actual publication of the statistical handbook.

The handbook should be published in a separate series such as that formed by the consecutive issues of the *Jahrbuch der Deutschen Bibliotheken*. The Office of Education might be in a position to include it as a subseries of its *Bulletin*, if federal funds were available. Preferably, however, it should be published by the A.L.A. because the editing must be done by librarians and because it will probably ultimately turn out that the only means of financial support will be contributions from libraries or librarians. Possibly the financing of the publication might be secured by making its purchase compulsory with institutional membership in the A.L.A. It is much easier to justify expenditures for collecting, compiling, and publishing statistical records to librarians than to legislators.

Most important, however, is to have one central office for the collection and publication of statistical data, thus avoiding duplication of work and making all statistical records available at one source. Preferably, that office should be somewhere within the framework of the A.L.A. and should include on its staff at least one individual from each type of library for which statistical records are collected. From the standpoint of library administrators, a central source would be particularly welcome, for it would eliminate the necessity of answering a dozen or more questionnaires each year.\(^6^6\) The figures on university libraries published for many years by James Thayer Gerould and, presently, by Lawrence Heyl, of Princeton; the figures on southeastern college and university libraries, published by McMillen; the few items of statistical information published in the *American Library Directory*; and many statistical publications of state library agencies and even of the Office of Education, would be rendered unnecessary by an exhaustive annual statistical handbook.

**Handbook's Inclusions**

The handbook should include at least as many libraries as the Office of Education compilations and at least as many categories on each of them as the A.L.A. tables. The peculiar advantages of the Princeton and Louisiana State figures, in presenting a particular type of library or geographical sector, could easily be duplicated by appropriate organization of information at the central statistical office. However, the advice of those who have had prior experience in compiling statistical data on libraries should be fully utilized by placing them on the committees suggested above.

If a statistical annual is undertaken, it should be done with a view to publishing it promptly and permanently. For all their other faults, the one great virtue of the A.L.A. statistics is that they can show almost two decades of continuity in their present form.\(^6^7\) They also are published promptly every year within about eight months of the end of the fiscal year, for most of the institutions they cover. The present wide distribution of the A.L.A. statistics could almost be duplicated if all

---

\(^{66}\) Furthermore, expansion of statistical categories published at present would obviate the necessity for the circulation of many of the questionnaires used by graduate students in compiling data for masters and doctors theses.

\(^{67}\) Burgess, *op. cit.*, p. 110-11, has compiled a table showing the growth of A.L.A. statistics of college and university libraries, by tabulating the statistical categories included each year since 1907.
institutional members of the A.L.A. received copies of the statistical handbook. If wider distribution is deemed necessary, it would not be too serious a duplication of effort and space to publish pertinent abstracts in the *A.L.A. Bulletin* and in *College and Research Libraries*.

Ultimately, the greatest hazard to any publication of the nature of the proposed statistical handbook is the danger of petrifying or falling into the control of individuals who are unwilling to subject it to alterations as changed conditions may require. Our statistical handbook must be capable of any addition or deletion, and there must be appropriate administrative organization for expediting and facilitating such changes whenever they are suggested and deemed worthy of adoption. Its contents and arrangement should be carefully reviewed every year or so by the higher control committee.

Above all, constant study of the problem of statistical records is necessary. We need specialized studies of what statistical records are necessary for each library department and for each type of library—in classifications even more detailed than college and university, small and large public, and school libraries. We need studies of the type of library statistical information needed for specific regions. We need particularly some group which can set up minimum standards for statistical content of annual reports of libraries. If this form of literature, which is so basic for all research studies in librarianship, is subjected to detailed and continuous investigation, the results may easily be far out of proportion to the effort expended.

---

**Cutter Classification**

*Because* of recurring requests for revision of the Cutter Expansive Classification, the A.L.A. Division of Cataloging and Classification has appointed a committee to make a thorough investigation of the need for such a revision. The first step is to get the names of all libraries, of every kind and size in the United States and Canada, using this classification or a modification of it. It would also be helpful to know what libraries have changed from Cutter to some other classification and how recently. If your library falls in either class, or if you can give any information that might be helpful, please address Dora Pearson, Cutter Classification Committee, Public Library, Washington 1, D.C.

*September, 1945*