



Measurement and Evaluation

RUTHERFORD D. ROGERS

AN AMAZING AMOUNT of ink has been consumed in presenting the case against conventional methods of measuring and evaluating public services, but there is as yet no evidence that the pen is mighty enough to penetrate the divinity that hedges our traditional approach to circulation and reference statistics. If we could plead ignorance of the follies we commit and the fallacies we perpetuate in this regard, we would be entitled to a modicum of sympathy. However, most of us resort to statistics as the best method of impressing public officials, corporation executives, and school, college, and university administrators in the full knowledge that this technique will not bear close scrutiny. Some of us have contemplated without too much relish that some perceptive recipient of our reports would some day ask the embarrassing questions that we have already asked ourselves about the significance of circulation and reference statistics.

It has been authoritatively stated that statistics on book use are found in the earliest available records;¹ however, it may be more significant that compilations of such statistics from a number of American libraries began in 1851.² The cases are indeed few in which present day annual reports fail to include figures on book use,³ despite the fact that the inherent weaknesses of such statistics are widely admitted by librarians in all kinds of libraries and even in publications put out by our professional associations.^{1, 4-16} The candor with which shortcomings are proclaimed on the one hand, and the frequency with which the continuation of present practices is advocated on the other, surpasses belief. The typical argument proceeds somewhat as follows: circulation statistics fail to present an accurate picture of what a library does; they have serious shortcomings; however, they have been used for a long time and are widely accepted; therefore, etc., etc.

At this point the reader has a right to ask for a bill of particulars against circulation statistics. Such a bill might read as follows:

Mr. Rogers is Chief of the Personnel Office of the New York Public Library.

1. Circulation statistics are a crude, quantitative measure, largely lacking in qualitative factors. To cite an extreme example, it is not proper to compare the circulation of a library which traffics exclusively in light romances, mysteries, and westerns with the circulation of a library composed of carefully developed subject collections and a critically selected fiction collection. The circulation of one technical book is more important than ten circulations of a popular novel, so the argument goes, although this unsupported statement cannot be categorically defended. In the same vein, it is not reasonable to compare a 1952 circulation which is 60 per cent nonfiction with a much larger 1932 circulation which was 65 per cent fiction. Furthermore, circulation statistics fail to reveal the significance of what is read, the impact of books upon the individual or upon society, or even whether or not the books circulated were read. Pursuing this line of reasoning, Hannah Logasa once wrote: "The circulation statistics in libraries tell the number of physical books furnished by the library. But the physical book is only paper, print and binding until it is reconstructed by a reader. By that standard, many books circulated are never born, because the reader has not read himself into the book. The only book circulation that really counts is one in which the reader has caused the book to be born." ⁶ Therefore, it is not enough to know that so many medical books, for example, were circulated. It is important to know whether or not the person who borrowed a given medical book was a biochemist, a physician, a medical student, a hypochondriac, an intelligent layman, or a curious young person, and to know to what use, if any, the information contained in the book was put.

2. The second item in the bill of particulars rests on the belief that book-use statistics, at best, are fragmentary. There are problems of unrecorded use in the case of open libraries, in stacks, in departmental collections, and in dormitory libraries, to cite but a few examples.

3. From the standpoint of comprehensiveness and comparability, statistics are greatly affected by a series of miscellaneous factors, and therefore it is not really possible to say that one library is doing a better job than another on the basis of circulation. These factors include the amount of money available for the purchase of duplicates, the presence or absence of a rental collection, curricular differences in schools and colleges, reserve book rules and the method of recording reserve book loans, the social structure and educational level of the community served by a public library, and even circulation rules and general operating procedures.

Measurement and Evaluation

As has been indicated, many librarians who grant the validity of all the foregoing criticism have been reluctant to turn their backs on the blandishments of statistics. These people have sought to perform a minor operation here and a major one there in order to overcome some of the weaknesses. Leon Carnovsky suggests that librarians "weight" books according to merit. For example, one circulation of Lippmann's *Good Society* would be "equivalent to half-a-dozen issues of Temple Bailey, with, say, a biography by Ludwig coming somewhere between. Lippmann would be scored 6; Ludwig, 3; and Miss Bailey, 1. It would be possible to go even farther, and assign varying values to fiction of varying literary merit."¹ This proposal is susceptible to criticism on the grounds of impracticability. Furthermore, such a plan would not give consideration to the reader and his purpose in reading a book. A social historian or literary critic might read extensively in light novels of the 1920's and thereby achieve a contribution to knowledge without the qualitative significance of the work being reflected statistically.

In attempting to refine circulation statistics, David Jolly⁸ has suggested ranking books according to "intrinsic worth," determining the age, sex, occupation, education, interest, and hobbies of each borrower and ascertaining "how much of the information is assimilated." H. I. Muller⁹ has taken a startlingly opportunistic attitude toward circulation statistics. He suggests, in effect, that the librarian determine those factors in the circulation statistics which bring good repute to the library, and then he proposes that we emphasize these factors in ingenious ways. For example, if it is more respectable to circulate a high percentage of nonfiction, one should not simply relate nonfiction statistics to total circulation but should show how many readers took some nonfiction, a much higher and more impressive figure.

An all-time low in pronouncements on circulation statistics was made in the October 1953 *A.L.A. Bulletin*. This article deplors the fact that libraries have not achieved, and are not likely to achieve, the per capita circulation of the depression years, despite the fact, according to the author's not-too-clear statistics, that we are getting much more financial support in 1939 dollars than we were in 1933. Reference is made to the "many new and additional services that take an important share of the library's income," but on an empirical basis it is concluded that this does not account for our increased well-being. The author then goes on to state:

In other words, the individual library is now lending about half

as many books per capita, per dollar of library income, per volume accessioned, per employee, or per anything else, as it did in the early 'thirties. . . . The statistics are indeed convincing, but what public librarian is ready to concede that his or her library can never expect to do as much "business" as was done twenty long years ago—to say nothing of doing as much per capita or per income dollar? No, librarians will continue working to get back up above that 1932–33 circulation line, realizing full well that circulation figures don't tell the full story but determined, nevertheless, that they aren't going to work forever in the shadow of a time "when we were *really* busy around here." ¹⁷

Conflicting statements in the above quotation to the effect that "statistics are indeed convincing" and "circulation figures don't tell the full story . . . nevertheless," are all too characteristic of many treatments of this subject. A parallel to the reasoning contained in the article under consideration may be found in comparing cost per-book-circulated in certain sub-branches which are purely mass circulation outlets open a few hours weekly, and cost per-book-circulated in branch libraries providing reference service, telephone service, leisure reading accommodations, story hours, community programs, and readers' advisory service on a forty to fifty hours per week basis. Certainly such comparisons are extremely hazardous just as a comparison between 1932 and 1952 is hazardous. This is further suggested by a letter to the editor published in the December 1953 *A.L.A. Bulletin*, commenting on the aforementioned article: "It seems to me there is a great fallacy in taking circulation statistics as a measure, *without taking into account many other factors*. . . . [During the depression] patrons were reading mysteries and westerns and any other kind of escape reading available, simply to forget their problems. . . . The fiction percentage in 1931 was 61%, in 1932, 62%, as against 27% in 1951 and 26% in 1952. . . . Twelve westerns a week count heavily against one technical handbook in a circulation record." ¹⁰

Before leaving the subject of circulation statistics, attention should be called to the Index of Public Library Circulation.¹⁸ If the criticisms advanced thus far have any foundation in fact or reason, the Index should be re-examined in terms of the philosophy which motivates its publication and possible use. It seems painfully apparent that the Index is the product of an almost purely statistical mode of thought, untrammelled by deeper professional implications discussed here.

Let us now shift our attention to reference statistics. Public librarians took the lead in introducing reference services in the third quar-

Measurement and Evaluation

ter of the nineteenth century, partly because the custodial concept of librarianship delayed the development in colleges, and partly because public librarians felt a need to justify expenditure of city funds by offering additional services.¹⁹ Most of the people who have considered the problem of reference statistics have concluded with Fremont Rider that "the finest service that every library gives is the very one that can never be measured."²⁰ Although this has led to the comforting thought that most important cultural and intellectual achievements cannot be "weighed on a scale, measured in inches, or price-marked in dollars and cents,"²¹ it has not noticeably decreased the use of statistics of reference service. However, increasing numbers of people recognize that "qualitative measurement" of reference service is a contradiction in terms,¹⁵ and they prefer the use of such terms as "evaluate," "interpret," and "judge."²¹

The small amount of space devoted to evaluation of reference service in some of our outstanding professional literature is surprising. The *Public Library Inquiry, A National Plan for Public Library Service*, and *The Administration of the American Public Library* acknowledge that there is a problem but largely disregard it.^{13, 14, 22} For the college library field G. R. Lyle²³ has made an excellent analysis of the subject. J. I. Wyer²⁴ dismisses the problem by saying that "It is doubtful whether statistics of this kind are worth the time and trouble involved in their preparation." *Post-War Standards for Public Libraries* is to be commended for its approach to the problem of "quality" in standards, despite some perplexing statements which seem to contradict each other. Many standards are stated in nonstatistical terms. For example, to meet responsibilities in the field of reference, larger libraries are urged to provide a card catalog, research materials in fields of community interest, instruction in the use of the library, and at least one trained librarian to do reference work during public service hours. However, one also learns that "Statistical records of reference and reading aid questions should be kept" and "the statistical enumerations of library service which are used as the bases for quantitative standards are generally both accurate and uniform in definition" despite the fact that "Library service embraces many intangible elements of quality and excellence which cannot be precisely measured" and the "definition and recording of reference and reading aid questions have not been standardized satisfactorily in public library practice."²⁵

Let us consider for a moment the methods that are generally used to "measure" reference service. Perhaps the most common is the keeping of a running tally of questions asked. This may be supplemented

by a classified list of questions, number of "search" and "research" questions as determined by certain time-to-answer categories, number of readers using service and time spent in library, record of actual questions asked, separate record of questions not answered, types of materials used in answering questions, number of bibliographies compiled, bulletins prepared, stack permits issued, telephone calls, letter requests, indexes made, club programs prepared, and interlibrary loan requests. This array is so formidable that it would be immediately suspect in the eyes of a competent administrator, and if the full battery of statistics should be turned upon public officials or their counterparts in school, college, and special library situations, these officials would be (a) completely confused, (b) impressed by the sheer weight of numbers, or (c) suspicious of librarians.

What are the criticisms of statistics of reference service other than mere proliferation?

1. Such statistics are not inclusive. Reference librarians say that only 60 per cent of the questions are recorded on a busy day,²¹ and an even smaller percentage of reference books used is ever recorded.^{26, 27}

2. Qualitative measurement, so-called, is unattainable although the essence of reference service is quality.^{15, 28} In response to a suggestion that the importance of a question might be determined by the place of the inquirer in the business, social, or faculty hierarchy, it has been humorously suggested that "a question from the [college] President might be rated as worth ten, a question from a Dean as seven, and so on down to a question from an undergraduate which ought to be worth minus one which . . . reduces the whole thing to the absurdity which it is."¹⁵

3. The effort to evaluate questions as to "search" and "research" has led to reliance on time differentials as a basis of identification. For example, a question that can be answered in fifteen minutes or less is a "search" question; a question requiring more than fifteen minutes is a "research" question. Obviously ridiculous though this technique is, it was once used by A.L.A. and the U.S. Office of Education.²⁸ As Mary N. Barton has pointed out, time distinctions are invalid because they do not give weight to the competence of the librarian; a question that requires two hours to answer receives the same value as one that takes sixteen minutes; and if a file of difficult questions and answers is kept, the first time a question is answered, it may be classified as "research," the second and succeeding times as "search."²⁷

Measurement and Evaluation

Nine out of ten of the people who have analyzed the flaws in reference statistics have not even attempted a solution for the problem of interpreting reference services to the public, but they have demonstrated that certain records can be of great value in internal management of the reference department. During the time that he was at the Grosvenor Library, R. W. Christ made substantial contributions to the management aspect of reference department records.²¹ He demonstrated that a complete record of questions and source of answers provided the following valuable information to the departmental supervisor: (1) those subordinates in need of training and closer supervision, (2) weaknesses in the reference and general book collection as well as important available materials that were being neglected, (3) the need for more adequate graphic material to help readers use the card catalog, and (4) changed techniques for answering questions from reference books, documents, and from the general book collection. H. C. Bauer also pointed out that a running record of reference work *not* done is of great value in self-evaluation of staff and collection.²⁹

Before attempting to reach any conclusions in regard to the use or misuse of public service statistics, let us consider possible reasons for our strong attachment to such statistics:

1. We have seen "science" and "research" lend respectability to related fields, such as education, and we hope that if we do likewise some of this respectability will rub off on us. Statistical hypnosis, to coin a phrase, has become an occupational disease. As C. E. Sherman¹² points out, ". . . librarians as a whole were, and have continued to be, influenced by mathematical considerations. Anyone who doubts this need only quiz the next half-dozen librarians he meets as to their understanding of the A.L.A. standards for public libraries."

2. Another occupational hazard, closely related to the preceding one, is the "Ph.D. approach." An anonymous colleague has observed that "the tendency to mechanize everything and measure everything by a centimeter rule is rapidly placing us librarians in the same class with Ph.D.'s from Columbia's Teachers College who have learned a technique of using standard measurements which they apply with so little imagination and cultural background that usually the result of their measuring is what the well-educated person knew intuitively all of the time."¹⁵

3. Sherman, who, with Lyle, has done as much clear thinking on this problem as anyone, says that "the mounting statistics of circula-

tion fit perfectly into the American's affection for big figures." ¹²

4. Number 3, above, is related to the Big Business approach to administration. Taking a leaf from the special librarian's notebook, Bauer ²⁹ says: "Management today expects results and demands statistical evidence of efficiency and accomplishment." J. A. Lowe ³⁰ echoes this sentiment when he says that the librarian who possesses statistics "approximates the position of a business executive who has definite elements of control, as standards against which he can check the actual efficiency of the business."

5. Even interdepartmental jealousy has contributed to the development of certain types of statistics. Margie M. Helm admits that exact measurement of quality in reference service is impossible, but she perceives the need for "objective data by which we could interpret the reference department to laymen and to college administrators. The circulation department has had its statistics of use. It seems that the reference department should devise some quantitative criteria by which its work could be interpreted or measured roughly. . . . If we can obtain objective data about . . . [reference] service it ought to be a better indication of the scholarly use of the library than circulation statistics." ²⁸ This is not a lone opinion as Edith Guerrier ³¹ confirms: ". . . the statistics gained by such a survey [of reference service] are of use in convincing library boards that the library's efficiency cannot be judged solely by circulation figures." In other words we create poor statistics to compete with bad ones.

6. Miscellaneous reasons for the use of statistics are candidly described by a British librarian: "It is pathetic to see, when reading a large number of annual reports, the almost hypnotic effect of statistics upon ourselves. How we give ourselves away. Some of us glory in them and find exhilarating evidence in every group of figures, for the wonderful services we are giving. Others introduce them apologetically and make a pretence of hiding them away but even so, how often they refer innocuously to some figure which purports to enhance their service. Regretfully too, one cannot fail to notice at times disparaging implications with respect to other library services, the more to enhance one's own."⁶

Most of this paper has been devoted to the shortcomings of circulation and reference statistics, but there is no desire to deprecate the utility of statistics in internal management situations where they can be employed with caution and the attention to nonstatistical factors known to good administrators. Progress in sampling techniques and

Measurement and Evaluation

the use of operational statistics in other fields encourage one to hope that precise, detailed, and expert analysis of small samples of circulation and reference work might produce valid conclusions of significance for release to appropriating bodies and patrons. However, if successful, these techniques will probably only be usable by large libraries, while the statistical problem, like so many other library problems, is not so restricted. It would not seem amiss, therefore, to look outside the statistical realm for a general solution to this problem.

With this in mind it is relevant to inquire into our objectives in using statistics. The ultimate purpose is almost invariably to ensure a certain level of financial support or to increase such support. This is true whether we present public service statistics to those who control the purse strings directly or whether we appeal to a wider audience which may indirectly influence appropriations. Since this is clearly a case of "public relations" with a motive, we might profitably examine a definition of the term: "Public relations have to do with the development and maintenance *by any legitimate means* of favorable attitudes on the part of the people with whom an agency comes into contact."³² The italics in the quotation are added by this writer and are used to emphasize the fact that public relations must be based on truth not propaganda.

It has been the author's good fortune to be associated with college, special, and public libraries, several with a reputation for excellence in collections and service which no proliferation of public service statistics could enhance or deprecate. It can be conversely reasoned that all the statistical finagling in the world cannot measurably improve the reputation of a library which has a poor collection and a surly or incompetent staff. We need day-to-day awareness on the part of our "community," whether it be a city, college, school, or corporation, of the services we have to offer. This means consistent and high-level public relations based upon the only sound foundation of any such program—competent and courteous service.

Bauer²⁹ confirms this point of view for the special library by saying: "In the final analysis, the true measure of a special library's service is the reputation gradually won for accurate, thorough, interested and intelligent help." Lyle has stated the case for college libraries in the following terms:

It is generally agreed that the functions of college libraries are not fully understood by those who are responsible for their welfare at the top level or by those who use its services. . . . Ask any student

what a librarian does and his generally unfurrowed brow immediately clouds up. The professor will parry the question with a perishable witticism or else state frankly that the librarian hands out the books at the loan desk. Yet no one should leap too quickly to the wrong conclusion. Students, faculty, and administrators know the importance of books in education and associate libraries and librarians with books. If they do not know what librarians do to make these books available . . . it is largely because they have not heard enough about such matters. . . . It seems reasonable to conclude, therefore, that efforts to improve public relations directly rather than as a mere by-product of good service are worth while.³³

Back in 1876 Samuel Green of the Worcester Free Public Library understood some of the elementary truths of library service which too many of us have forgotten: "The more freely a librarian mingles with readers, and the greater the amount of assistance he renders them, the more intense does the conviction of citizens, also, become, that the library is a useful institution, and the more willing do they grow to grant money in larger and larger sums to be used in buying books and employing additional assistants."³⁴

References

1. Carnovsky, Leon: Measurements in Library Service. In Joeckel, C. B., ed.: *Current Issues in Library Administration*. (University of Chicago Studies in Library Science) Chicago, University of Chicago Press, 1939, pp. 240-263.
2. Thompson, L. S.: History of the Measurement of Library Service. *Library Quarterly*, 21:94-106, April 1951.
3. Lyle, G. R., et al.: *Administration of the College Library*. 2d ed., rev. New York, H. W. Wilson, 1949, p. 508.
4. *Ibid.*, p. 536.
5. Logasa, Hannah: Figures Tell But Half the Story. *Library Journal*, 72: 999-1000, July 1947.
6. McClellan, A. W.: Reflections on the Use of Statistics in Public Libraries. *Library Association Record*, 52:182-186, June 1950.
7. Wilson, L. R., and Tauber, M. F.: *The University Library; Its Organization, Administration, and Functions*. (University of Chicago Studies in Library Science) Chicago, University of Chicago Press, 1945, p. 515.
8. Jolly, David: Evaluation of Library Service. *Wilson Library Bulletin*, 14: 296-299, Dec. 1939.
9. Muller, H. I.: Varieties of Presenting Fiction and Non-fiction Circulation in Libraries. *Library Journal*, 66:700-703, Sept. 1, 1941.
10. Shelly, Adah: Letter to the editor. *A.L.A. Bulletin*, 47:509, Dec. 1953.
11. Haygood, W. C.: *Who Uses the Public Library*. (University of Chicago Studies in Library Science) Chicago, University of Chicago Press, 1938, p. 4.
12. Sherman, C. E.: The Definition of Library Objectives. In Joeckel, *op. cit.*, ref. 1, pp. 22-46.

Measurement and Evaluation

13. American Library Association. Committee of Post-War Planning: *A National Plan for Public Library Service*. (Planning for Libraries No. 3) Chicago, The Association, 1948, p. 145.
14. McDiarmid, E. W., and McDiarmid, John: *The Administration of the American Public Library*. (Illinois Contributions to Librarianship No. 3) Chicago, American Library Association and the University of Illinois Press, 1943, p. 229.
15. Henry, E. A.: Judging Reference Service. *Library Journal*, 64:358-359, May 1, 1939.
16. Lowe, J. A.: *Public Library Administration*. Chicago, American Library Association, 1928, p. 143.
17. Alvarez, R. S.: Public Library Circulation Still in the Red. *A.L.A. Bulletin*, 47:435-436, Oct. 1953.
18. Goldhor, Herbert: Index of American Public Library Circulation. *A.L.A. Bulletin*, 43:334-335, Nov. 1949.
19. Rothstein, Samuel: The Development of the Concept of Reference Service in American Libraries, 1850-1900. *Library Quarterly*, 23:1-15, Jan. 1953.
20. Rider, Fremont: Library Cost Accounting. *Library Quarterly*, 6:331-381, Oct. 1936.
21. Christ, R. W.: Recording Reference Service. *College and Research Libraries*, 8:23-27, Jan. 1947.
22. Pierce, W. O'D.: *Work Measurement in Public Libraries*. New York, Social Science Research Council, 1949, p. 186.
23. Lyle, *op. cit.*, p. 588.
24. Wyer, J. I.: *Reference Work*. (Library Curriculum Studies) Chicago, American Library Association, 1930, p. 250.
25. American Library Association. Committee on Post-War Planning: *Post-War Standards for Public Libraries*. (Planning for Libraries No. 1) Chicago, The Association, 1943, pp. 12-13, 16, 27.
26. Lowe, *op. cit.*, p. 153.
27. Barton, Mary N.: Administrative Problems in Reference Work. In Butler, Pierce, ed.: *The Reference Function of the Library*. (University of Chicago Studies in Library Science) Chicago, University of Chicago Press, 1943, pp. 218-248.
28. Helm, Margie M.: Interpreting Reference Service. *A.L.A. Bulletin*, 34:115-116, Feb. 1940.
29. Bauer, H. C.: Evaluation of Service in a Technical Library—Library Statistics. *Special Libraries*, 30:296-297, Nov. 1939.
30. Lowe, *op. cit.*, p. 145.
31. Guerrier, Edith: The Measurement of Reference Service. *Library Journal*, 61:529-531, July 1936.
32. Civil Service Assembly of the United States and Canada: *Public Relations of the Public Personnel Agencies*. Chicago, The Assembly, 1941, pp. 5-6.
33. Lyle, *op. cit.*, p. 488.
34. Green, Samuel as quoted in Rothstein, *op. cit.*