But What Does the Data Mean? Getting From What Happened to Why it Happened

Sharon L. Baker

During the 1960s and early 1970s, the American economy was so favorable that funding for all types of libraries increased. Librarians received most of the resources they needed to implement or maintain services even though they collected few data on the real success of library programs. In the last fifteen years, however, the average cost of running a library has risen faster than its income. Today, funding organizations expect libraries to continue providing quality services while keeping costs down. They also want "proof" that library programs are operating efficiently and effectively.

These changes in the funding climate and the spread of sophisticated evaluation techniques through society in general have led various state and national library associations to promote the use of performance measures in all types of libraries. Some libraries have been slow to adopt these measures,[1] but their use is growing.

The Macroevaluation of Library Services: Learning What Happened

Such performance measures generally emphasize the *macroevaluation* of library services. As Baker and Lancaster (1990) explain in some detail, macroevaluation studies measure the success rate of a system; that is, they describe *how well* it operates. The results of macroevaluation studies can usually be expressed in quantitative terms, such as the percentage of reference questions answered accurately. For example, and as Figure 1 shows, the twelve measures discussed in *Output Measures for Public Libraries*[2] are all macroevaluation measures. Because such measures show the level of performance at which a service is operating at a specific date, they serve as a benchmark. Library directors can use performance data from their own libraries and from comparable libraries to support the argument that more resources are needed to improve program quality. Then, if resources are subsequently added, the library director can compare results with this benchmark to see if the service has improved. Benchmark figures can also be reviewed to determine if the quality or quantity of service is declining. Indeed, many libraries collect such performance data to serve as an early warning signal for trouble spots.[3]

Unfortunately, librarians who collect this type of benchmark performance data still have a major problem. While they know what happened in regard to a given library program, they often do not know *why* a program is or is not successful. That is, the macroevaluation measures collected do not give librarians enough information to make intelligent changes to improve service quality. This may explain Schlachter and Belli's discovery that seventy-eight percent of the California public libraries that collected performance data made no changes based on the findings.[4] Some needed changes may not have been made for quite valid reasons, such as a lack of immediate resources to solve specific problems. But the fact that so many libraries failed to make *any* changes may indicate that collection of this type of macroevaluation-oriented performance data does not, in and of

FIGURE 1.

Macroevaluation Measures Appearing in *Output Measures for Public Libraries* (Van House et al., 1987)

Annual library visits per capita
Registration as a percentage of population
Circulation per capita
In-library materials use per capita
Turnover rate
Title fill rate
Subject and author fill rate
Browsers' fill rate
Document delivery rate
Reference transactions per capita
Reference completion rate
Program attendance per capita

Sharon L. Baker is an assistant professor at the University of Iowa, School of Library and Information Science in Iowa City, Iowa.

itself, provide enough useful information to improve services.

The Microevaluation of Library Services: Learning Why it Happened

Diagnostic information which *can* be used for improvement comes from *microevaluation* of library services. Microevaluation investigates *how* a system operates and *why* it operates at a particular level — that is, what makes it work well or badly. The most important element of this diagnosis is identifying reasons for particular failures. For example, while it is nice to know that fifty percent of a library's patrons did not receive complete and accurate answers to their reference questions, improvements cannot really be made unless the causes of the problem are pinpointed. A microevaluation would examine whether the reference librarians failed to verify the users' "real" information needs, used poor strategies to search the catalog or other bibliographic tools for the answers, or were too busy to accompany patrons to the shelves to show them the specific items that could answer their questions. Microevaluation would also look at other reasons for failure, such as collection inadequacy or poor subject access in the card catalog. This type of microevaluation study is of greater practical use to the librarian because it provides guidance about which actions might be taken to improve reference accuracy. That is, a microevaluation study of this nature tells us what the performance measure (the fifty percent accuracy rate) really means.

Microevaluation investigates *how* a system operates and *why* it operates at a particular level . . .

Although most of the performance measures promoted by library associations are examples of *macroevaluation*, librarians can fairly easily expand their data collection efforts to determine how and why these success rates were obtained — that is, to include *microevaluation*. Let's take a simple example.

Output Measures for Public Libraries suggests that one performance measure — title fill rate — be collected using a simple patron questionnaire.[5] Generally, each patron who enters the library during a selected week is given a questionnaire on which to indicate the works being sought and whether or not they are found. The form is turned in as the patron leaves the library. At the end of the week, the total number of titles found by patrons is divided by the total number of titles sought. This gives the library's overall success rate in filling patron requests for specific items.

Evaluators who stop here will know what is happening — that is, what proportion of a patron's needs for specific materials have been met — but they will not have the diagnostic information necessary to increase their fill rates in the future. They must go beyond such macroevaluation studies and determine if the reasons *why* particular titles are unavailable fall into perceivable patterns. For example, are there major collection gaps? Are popular titles owned but in quantities insufficient to meet patron demands? Are purchased titles stalled somewhere in technical processing so patrons still do not have access to them? Are the reshelving procedures so slow that books are sometimes present in the library but unshelved so that patrons cannot find them?

Such a microevaluation study is actually quite easy to perform, if the evaluator simply carries the data collection efforts a bit further. In the above study of fill rate, the evaluator should not stop at simply asking patrons to indicate on a questionnaire whether they found the titles they were seeking. Rather, as each questionnaire is turned in, the evaluator should check the catalog, the shelves, and the circulation area to determine why the patron failed to find desired items. For example, several major problems might inhibit patron access to specific items: acquisitions barriers, circulation interference, patron errors (in using the catalog or in searching for materials in the stacks), or other library errors like misshelving. As Figure 2 shows, each of these problem areas can be broken down even further. In fact, the finer the analysis, the more likely the evaluator is to figure out why books in this library are unavailable for use when patrons want them. Once the evaluator has determined the problems that occur most frequently, library practice can be changed to prevent, or at least decrease, the chances of those problems recurring. For instance, if many titles are unavailable because they are checked out to other patrons, the library can either shorten the length of the loan period for popular titles or can buy more copies of them.

Virtually all the macroevaluation measures that are recommended by library associations (measures of fill rate, reference accuracy, speed in interlibrary loan or document delivery, etc.) can be used as the first step in a microevaluation study. In most libraries, a committee of profes-

FIGURE 2.

List of Reasons for Nonavailability of Titles Developed by the Iowa City (Iowa) Public Library

The Acquisitions Barrier

The library does not own the title.

The library has ordered it, but it has not yet been received.

The library has received the title, but it has not yet been cataloged and processed.

The patron does not know about other options such as requesting that the title be purchased or asking for it to be obtained through interlibrary loan.

Circulation Interference

The item is checked out to another borrower.

The item is checked out to Technical Services to be repaired, re-bound, re-cataloged, or re-labeled.

The item is checked out to "Missing" and has not yet been replaced.

The item is long overdue from another library borrower and no decision has been made about whether to replace it.

Library Error

The item is checked in but is not yet reshelved.

The item is mis-shelved.

The call numbers on the item and in the catalog do not agree.

The library is unaware that the item is missing (e.g., it has been stolen).

The item was not properly checked in; the catalog indicates it is checked out, but it is on the shelf.

The item is currently in use by a staff member but it is not checked out.

User Error

The user cannot find the item in the catalog (e.g., due to incorrect title or author information or incorrect search techniques).

The user finds the bibliographic record in the catalog, but misinterprets the information. For example, he assumes that the bar code or publication date is the call number or he records the call number incompletely or in the wrong number order.

The user locates the title in the catalog and copies the correct call number down, but he cannot find the location. For example, he doesn't understand the significance of certain terms or symbols in the call number; he can't find the location referred to; he makes mistakes in the alphabetical or Dewey order; he doesn't understand the sequence of shelving units.

The user does not ask a staff member for help at the catalog or at the shelf. This could be due to his not being able to find or identify a staff member, to his finding a staff member already occupied with other patrons, or to his fear of asking a staff member for help.

sional librarians can examine any type of raw performance data, isolate those services with inadequate performance levels, list a number of possible reasons why performance might be bad, and develop a "quick and dirty" study to see what is actually causing the problem. For example, if a library discovers that few interlibrary loans are filled within an acceptable period of time (say ten days), the librarians can generate a list of possible causes of the poor performance. These may be related to the characteristics of materials requested (such as the date and the form of publication), the size or training of the interlibrary loan staff, membership in a library network, or factors relating to other libraries (e.g., although materials are requested quickly, some requesting libraries may be slow to fill the orders). A fairly quick evaluation can identify which of these is the most likely reason for poor performance. Library staff can then work to reduce the problem. For example, if two libraries within an interlibrary loan network are found to be very slow in filling material requests, staff can be advised to seek materials from other libraries first.

Some librarians may feel that they lack the necessary expertise to conduct microevaluation studies. And indeed, issues of validity and reliability should be considered to ensure that the performance data is accurate.[6] Validity refers to whether the evaluator is actually measuring what is intended and to whether generalizations can be made from the data collected. Reliability refers to whether the evaluator can expect to obtain the same results if the data is collected at a later date or by a different evaluator.

Aid for librarians who need help with these or other methodological problems is available in several forms. Library schools, library associations, and state library agencies may provide consultation services, recommend consultants who are experts in evaluation, or present workshops on evaluation techniques. Librarians collecting performance data can also read any of several recently published books on the topic, such as *Measurement and Evaluation of Library Services*,[7] *If You Want to Evaluate Your Library . . .*,[8] and *Are We There Yet? Evaluating Library Collections, Reference Services, Programs, and Personnel*.[9] These books discuss some of the finer points of collecting performance data and cover both the macroevaluation and the microevaluation of library services. The books also recommend variations on particular themes (e.g., using separate fill rate studies for each branch of departmental library and for each format of materials owned). The Baker and Lancaster title summarizes findings of past evaluative studies as well.

Summary

Because of funding limitations facing libraries today, librarians are collecting more performance data. Unfortunately, most librarians limit the usefulness of the data by collecting information that focuses almost exclusively on what happened in a given situation. Such data is useful because it establishes a benchmark figure against which future data can be compared. In order to make real improvements in service, however, librarians also need to explore *why* and *how* things happen in libraries. That is, librarians need to determine the causes of particular problems so that effective changes can be made.

References

1. Sharon L. Baker, "The Use of Output Measures for Public Libraries in North Carolina Public Libraries." (Iowa City, Iowa: School of Libraries and Information Science, 1987) Eric document number ED 288 538.
2. Nancy A. Van House et al., *Output Measures for Public Libraries: A Manual of Standardized Procedures.* 2d ed. (Chicago, Ill: American Library Association, 1987).
3. Association of Research Libraries, Office of Management Studies, *Planning for Management Statistics in ARL Libraries.* Systems and Procedures Exchange Center, Kit 134. (Washington, D.C.: Association of Research Libraries, 1987).
4. Gail Schlachter and Donna Belli, "Program Evaluation — An Alternative to Divine Guidance." *California Librarian* 37 (October 1976): 26-31.
5. Van House.
6. Terry L. Weech, "Validity and Comparability of Public Library Data: A Commentary on the Output Measures for Public Libraries." *Public Library Quarterly* 8 (1988): 7-18.
7. Sharon L. Baker and F. W. Lancaster, *Measurement and Evaluation of Library Services.* 2d ed. (Washington, D.C.: Information Resources Press, in press).
8. F. W. Lancaster, *If You Want to Evaluate Your Library . . .* (Champaign, Ill.: University of Illinois, Graduate School of Library and Information Science, 1988).
9. Jane B. Robbins and Douglas Zweizig, *Are We there Yet? Evaluating Library Collections, Reference Services, Program and Personnel* (Madison, Wis: School of Library and Information Studies, University of Wisconsin-Madison, 1988).