
From Measurement to Management: Using Data Wisely for Planning and Decision-Making

STEVE HILLER AND JAMES SELF

ABSTRACT

The wise use of data, information, and knowledge in planning, decision-making, and management can improve library performance. While libraries have long collected data, however, it is only recently that they have begun to use it effectively in library management. This article provides an overview of data use in libraries, organizational barriers, and support issues, as well as examples of libraries that have successfully integrated data acquisition, analysis, and application into management.

INTRODUCTION

Data can be defined and used in a number of different ways. Our simple definition is that data are records of observations, facts, or information collected for reference or analysis. Data may take a number of forms, such as transactions, observations, surveys, or interviews. All of these provide data, that is, observations, both quantitative and qualitative, from which inferences may be drawn by means of analysis.

Libraries have long collected data about library operations, primarily inputs such as the size of collections and staff or expenditures. However, the degree to which they were used, or could be used, for decision-making in library management varied widely. Recently, there has been a voluminous increase in library-related data, not only in transactional information from online library systems and electronic resources usage but also from efforts to gain more direct user input through surveys, focus groups, and other methods. Funding and accrediting bodies are also asking libraries to demonstrate

their impact on the user community through performance measurements that are based on outcomes and data. While many libraries recognize the value of using data for planning and decision-making, they are unsure how to collect, analyze, and apply the data effectively in library management.

This concern is not new. Libraries have struggled for years with how to utilize statistics and other data to enhance library effectiveness. Nearly two decades ago, Allen posed these questions at an international conference:

The failure of library statistics to solve all the problems that library management would have them solve may not, however, be entirely the fault of the statistics. A number of questions may be reasonably asked. Do librarians collect the appropriate statistics? Are the statistics collected either accurate or comparable among similar libraries? Do we ask valid questions of the data? And above all, do we know how to manipulate and interpret statistical information? All too often the answer to these questions is "no." (Allen, 1985, p. 212)

Although many libraries have measured aspects of library activity or operations, why have the majority failed to use data effectively in management? What are the obstacles and barriers? Are there strategies and programs that have worked well, providing models from which we can learn? This article will review both the problems and successes involved in using data wisely in library management and decision-making.

TRADITIONAL USES OF DATA IN LIBRARIES

Libraries have generated and collected data related to their operations for many years. Statistical data in such areas as expenditures, number of books purchased, and staff size were gathered and reported to appropriate administrative bodies or groups. Gerould was among the first to discuss the practical value of comparative data:

No questions arise more frequently in the mind of the progressive librarian than these: Is this method the best? Is our practice, in this particular, adapted to secure the most effective administration? Are we up to the standard set by similar institutions of our class? These questions are of the most fundamental type, and upon the success with which we answer them depends much of the success of our administration. (Gerould, 1906, p. 761)

Gerould further elaborated on the statistical categories that would prove helpful in library administration and management. These included facilities, collections, finances, staff, salaries, ordering and processing, cataloging, collection use, reference transactions, and departmental libraries. He began collecting and publishing data in 1907–8 from a select group of academic research libraries, and the practice continued (after his retirement) until 1962, when the Association of Research Libraries (ARL) took over the collection, compilation, analysis, and distribution of statistics. While these early statistics provide an invaluable record documenting the historical development of

American academic research libraries, there is little evidence on how they were actually used to improve library management and decision-making. While it is likely that comparisons with other libraries may have added fuel to budget requests for increased funding, local statistics were more likely to be used for library planning. For example, the best data for projecting collection growth and the need for expanded facilities "are found in the individual library's statistical history" (Metcalf, 1986, p. 155). In his work on the Gerould statistics, Molyneux (1986) included library collection growth as the only example of how this data set could be used.

Comparative statistics were also used to develop standards, especially by library organizations. Such standards might specify the minimum number of volumes, staff, user seating, and other library measures. Efforts were also made to incorporate these standards or other statistical data into budget allocation, both at the institutional level and within the library. Library funding models or formulas such as Clapp-Jordan in the 1960s (and subsequent variants) endeavored to tie a recommended collection size to measures such as number of faculty, undergraduate students and majors, and graduate students at the masters and doctoral levels. Internal allocation models for collection development by subject area also used faculty and student numbers correlated to academic departments, as well as data related to publishing output, costs, type of materials, loans, and other use measures. While these were clearly efforts to use standards and data in library management, they were based on assumed linkages rather than research. Because these data were input centered, the link to outcomes were, at best, difficult to measure. As Clapp and Jordan admitted:

The formulas described in this article have been developed in an attempt to find a method for estimating the size for minimal adequacy of academic library collections more convincingly than can be done with existing criteria. It may be validly objected that little more has been accomplished than to transfer the locus of conviction from an unknown whole to the unknown parts, of which the whole is composed. (Clapp & Jordan, 1965, p. 380)

Of greater utility to libraries were local research studies that examined specific library services and processes undertaken in order to improve library performance. These included evaluating such activities as cataloging efficiency, card catalog use, reference services, collection use, interlibrary loan and document delivery, facilities and access, library systems, budgeting, and personnel. F. W. Lancaster's influential 1977 book, *The Measurement and Evaluation of Library Services*, provided the first systematic review of studies designed to measure and assess library performance. Lancaster also covered the different methods that could be used for evaluation. He made the important distinction between broad-based input/output data ("macroevaluation") and more focused analysis and interpretation of system processes ("microevaluation"):

Macroevaluation measures *how well* a system operates, and the results usually can be expressed in quantitative terms (e.g., percentage of success in satisfying requests for interlibrary loans). It reveals that a particular system operates at a particular level, but it does not, in itself, indicate why the system operates at this level or what might be done to improve performance in the future. Microevaluation, on the other hand, investigates *how* a system operates and *why* it operates at a particular level. Because it deals with factors affecting the performance of the system, microevaluation is necessary if the results of the investigation will, in some way, be used to improve performance. (Lancaster, 1977, p. 2)

In a subsequent paper Lancaster and McCutcheon went on to state,

Many of the studies conducted in the last ten years that can be grouped under the general heading of quantitative methods, are pure macroevaluation because they rarely go beyond producing data. In order to improve the service, we need microevaluation. . . . This type of analysis, although we use figures in our analysis, is more or less non-quantitative. It is interpretative. The investigator is very much concerned with using the figures acquired through quantitative procedures, to make reasonable decisions on what needs to be done to raise the level of performance. (Lancaster & McCutcheon, 1978, pp. 13–14)

Library Automation and Data Generation

The development and implementation of library-related systems for information retrieval, cataloging, and circulation coupled with the increased use of computers for quantitative analysis in social sciences helped move library education to a more systems-based approach in the late 1960s and 1970s. A new generation of library educators and librarians emerged who were equipped with quantitative skills and a structured social science approach to problem-solving that resembled Lancaster's microevaluation. Swisher and McClure addressed the need for "developing a research plan and analyzing data in such a way that practicing librarians can make better decisions and improve the overall effectiveness of their libraries" (Swisher & McClure, 1984, p. xiv). They called this type of applied activity "action-research" and defined it as the "ability to formulate questions about library services and operations, collect empirical data that appropriately describe factors related to those questions, and analyze those data in such a manner that summary descriptive information will be produced to answer the original question and implement actions/decisions to increase library effectiveness" (Swisher & McClure, 1984, p. 2).

By the early 1980s automated library systems could generate copious amounts of data and reports on circulation, cataloging volume, and use of catalogs and bibliographic databases. It was envisioned that these systems would form the core data elements of the emerging Management Information Systems (MIS) and Decision Support Systems (DSS) that would underpin good library management and decision-making in the future. Heim

defined an MIS as "A system that provides management with information to make decisions, evaluate alternatives, measure performance, and detect situations requiring corrective action" (Heim, 1983, p. 59).

Decision support systems were seen as supporting broader administrative and management decisions. Dowlin and McGrath envisioned this scenario in the not too distant future:

The goal for the DSS is for the library director or manager to use a terminal to ask the DSS: How is the library today? The system would respond with such comments as: "terrible," "lousy," "fair," "good," "not bad," or "great." The questioner could then ask why. The system would respond with a summary report of all of the indicators using predefined criteria that would indicate exceptions. (Dowlin & McGrath, 1983, p. 58)

Yet at the same conference in 1982 where Dowlin and McGrath presented their view of how systems data would be used in management (Library Automation as a Source of Management Information), Shank expressed his doubts:

The whole system seems to be put into place as a perpetual motion machine all too often installed without there being any analysis of what to do with the data. . . . It is not clear, what, if anything, can be done about whatever the data purports to show Data rejection occurs because there is a lack of understanding as to what data and information will sustain decisions about the value of services. (Shank, 1983, pp. 4–5)

Shank's comments certainly illustrated the need for the data to be analyzed, presented, and reported in a manner that could be easily understood and grasped by managers, administrators, staff, and other stakeholders. Burns noted in ARL Spec Kit 134 (*Planning for Management Statistics in ARL Libraries*):

The collection and use of management statistics is of almost universal concern to academic library administrators as part of their efforts to accurately describe their libraries' performance, evaluate and enhance effectiveness, and plan for the future. Although the need for management statistics and the potential for their use in decision-making is acknowledged by research libraries, most are still searching for ways to reconcile internal needs with external requirements, and to develop systems for effective use of statistics. (ARL, 1987, p. i)

Two years later, Vasi observed in ARL Spec Kit 153 (*Use of Management Statistics in ARL Libraries*) that, while many libraries gathered statistical data, there appeared to be little use of such data in planning and evaluation and a distinct lack of analysis:

Despite the wide range of possible uses for management statistics listed here, the predominant use for statistics is for comparison purposes—either with other institutions or year-to-year within libraries. It may be more valuable to ask why statistics are not used more frequently for other than comparative purposes. Comparative statistics seem to be

ends-in-themselves rather than as initial steps in an analysis of a library's operations or in quality of service. In almost all documents submitted, statistical reports were not accompanied by narrative analysis of the meaning of the data. . . . Why aren't more creative, analytical uses made of the large amount of statistics collected? Another phrasing of the question might ask how library managers use statistical data to make decisions on basic library goals and management of resources. (ARL, 1989a, p. ii)

Although automated systems made the process of generating process-related statistical data easier, whether these were the appropriate data and how to utilize them were still problematic for most libraries. No matter how well the MIS or DSS models looked in theory, they rarely worked in practice to meet the needs of administrators and managers. As Young and Peters summarized, "the appealing elegance, simplicity and effectiveness of MIS as an ideal has been difficult to design and implement in the real world." (Young and Peters, 2003, p. 1754). There were plenty of data, but they were not necessarily the right data and most library staff lacked the necessary interpretation, analysis, and presentation abilities to apply data effectively in management and decision-making.

McClure recognized these problems and pointed out several areas where more research was needed on MIS and DSS, including models and design considerations, hardware/software needs, and organizational impact and behavior. He posited a series of research questions, including, "What organizational climates and administrative assumptions facilitate the effective use of library MIS and DSS?" (McClure, 1984, p. 39). In examining the promise of microcomputer systems to improve management decision-making in libraries, McClure cautioned, "Regardless of the quality of and state-of-the-art of microcomputing hardware and software, the organizational context can preclude effective development of microcomputer-based decision making" (McClure, 1986a, p. 41).

Organizational issues rose to the forefront in McClure's study on the use of cost and performance measures by middle managers in ARL libraries. He concluded that, "The views and attitudes expressed during these interviews frequently suggest that significant organizational change will be necessary *before* cost and performance measurement data can be integrated successfully into academic library decision making" (McClure, 1986b, p. 329). McClure went on to recommend professional- and organizational-level strategies to increase the use of data in decision-making:

1. Review existing management styles and organizational climates within the academic library.
2. Increase the knowledge level of the importance and potential applications of cost and performance measurement data.
3. Develop administrative systems that support the identification, collec-

tion, organization, analysis, and reporting of cost and performance measure data.

4. Establish reward structures for librarians who use cost and performance measurement methodologies for library decision-making. (McClure, 1986b, pp. 332–333)

Organizational Development and Data Use

Organizational development concepts began to be incorporated into library management studies and reviews by the early 1970s. In particular, strategic planning, workplace environment, staff development, decentralized decision-making, and organizational efficiency were emphasized as critical components of a management review. The Management Review and Analysis Program (MRAP) sponsored by ARL helped bring organizational development into academic libraries. MRAP was an institutional self-study that saw strategic planning at the heart of organizational change. Data for decision-making played a key role in the organizational environment as it was used in each phase of the reiterative planning and action process. In organizational development, this was known as “action research” and defined by French and Bell as the following:

Action research is the process of systematically collecting research data about an ongoing system relative to some objective, goal or need of that system; feeding these data back into the system; taking actions by altering selected variables within the system based both on data and on hypotheses; and evaluating the results of actions by collecting more data. (French & Bell, 1999, p. 130)

The elements of the strategic planning process in libraries as it evolved during the 1980s included the development of a mission, vision, and values statement along with an environmental analysis that looked at both external and internal activities and trends. The formulation of goals and objectives as part of action planning, implementation, and evaluation followed. Data collection and utilization as part of this process became critical in two areas: assessing current library performance and measuring progress toward achievement of goals and objectives. Gardner noted in his introduction to an ARL Spec Kit 158 (*Strategic Plans in ARL Libraries*) the “importance of success measures and of the need for libraries to develop more ways of understanding its programmatic strengths and weaknesses” (ARL, 1989b, p. ii). In one of the first books published on strategic planning in libraries, Riggs wrote, “The importance of maintaining comprehensive statistics, conducting well-designed surveys, and using reliable performance measures cannot be overemphasized. Data derived from these records/studies will be crucial when the library’s goals and objectives are being scrutinized” (Riggs, 1984, p. 20).

Hernon and McClure noted that this type of formative evaluation takes

the greatest effort because it "requires the existence of clearly stated library goals and objectives" and the "establishment of regular data analysis and collection procedures" (Hernon & McClure, 1990, p. 10). They emphasized that political and organizational barriers must often be overcome for evaluative data to be used effectively.

As strategic planning took hold in many academic libraries, the need for developing performance measures closely linked to library goals, objectives, and action plans grew. While this article focuses primarily on academic libraries, a number of public libraries had been working with performance measures since the 1970s (see DeProspero, Altman, & Beasley, 1973). These pioneering studies established a framework for performance measures based on practical ways to measure library services, user success as a primary factor in service quality, and development of similar measures that could be employed across different libraries to provide comparative information. Public libraries by their nature are more involved in community analysis and use demographic and other related data to tailor services and collections to the local population. Public libraries also compete with other public agencies for local support from the governing body or directly from taxpayers. This is an added incentive to demonstrate the economic and social value of the library to the community using relevant data sources.

Van House and colleagues provided this justification for developing and using performance measures in academic libraries: "Carefully selected and intelligently used, output measures enable librarians to determine the need to which objectives are accomplished, set priorities for resource allocation, justify services, and demonstrate the degree of library effectiveness to the library's parent organization and other agencies." (Van House, Weil, & McClure, 1990, p. 13).

User-Centered Libraries and the Culture of Assessment

The concept of the user-centered library emerged in the late 1980s and early 1990s, fostered by strategic planning, total quality management, the external demands for accountability and measurable outcomes, and rapidly changing information and budgetary environments. Management strategies emphasized the need to focus on the customer and customer needs rather than organizational inputs and tasks. As Stoffle and her colleagues at Arizona stated:

Libraries must move from defining quality by the size of the inputs—and especially from valuing staff and collection size as "goods" in and of themselves. They get away from an internal professional evaluation of quality rooted in the context of what librarians agree that libraries do. All services and activities must be viewed through the eyes of the customers, letting customers determine quality by whether their needs have been satisfied. Librarians must be sure that their work, activities and tasks add value to the customer. (Stoffle, Renaud, & Veldof, 1996, pp. 220–221)

To accomplish this, user- or customer-centered libraries "collect data and use them as the basis for decision-making rather than rely on subjective impressions and opinions" (Stoffle, Renaud, & Veldof, 1996, p. 221). The keys to the success of the user-centered library can be found in understanding user needs, information seeking and using behaviors, user satisfaction, and providing the organizational focus and support on positive user outcomes.

Lakos, Phipps, and Wilson (2002) have promoted the concept of establishing a positive organizational climate for data-based decision-making through the development of a culture of assessment in libraries. Lakos's definition summarized their work:

A Culture of Assessment is an organizational environment in which decisions are based on facts, research and analysis, and where services are planned and delivered in ways which maximize positive outcomes and impacts for customers and stakeholders. A culture of assessment exists in organizations where staff care to know what results they produce and how those results relate to customers' expectations. Organization mission, values, structures, and systems support behavior that is performance and learning focused. (Lakos, 2002, p. 313)

Lakos has written extensively on the organizational components of a culture of assessment. He notes:

The prerequisites for a culture of assessment are supportive leadership, openness, integrity and trust. Developing positive assessment values and acceptance for assessment work is much easier in an organization where these prerequisites exist. Assessment is not about systems and tools, it is about people working together toward a common goal. (Lakos, 1999, p. 5)

According to Lakos, administrative leadership is critical: "The presence of visible leadership cannot be emphasized enough. Leadership is paramount for any organizational culture change to take hold, to be taught to the organization, and sustained over time until it becomes ingrained" (Lakos, 2002, p. 316).

It seemed as though all the building blocks for effective data use in management were in place by the end of the millennium. Library systems, microcomputer technology, and more recently the ubiquity of the Internet all helped provide increasingly powerful and easy to use tools for data collection and analysis. Spreadsheets and statistical analysis packages resided comfortably on desktop computers or local networks and were part of the library toolkit for measurement and evaluation. Organization development was firmly entrenched in many libraries with ongoing and iterative strategic planning, staff development, staff empowerment and reorganization, and a strong focus on quality and the user.

While some libraries had made the transition from measurement to informed use of data in management, the difficulties associated with using

data effectively remained for many libraries. Hiller (2002a) described the obstacles of organizational structure and inadequate leadership, librarian unease with quantitative analysis, lack of good data analysis and presentation skills, and the need to develop meaningful measures as major barriers to more extensive use of statistical data in libraries. The task of writing indicators and measuring performance turned out to be a complex activity with mixed results at best. Kyrillidou noted that "Performance indicators are being developed from data that can be easily gathered. Of course, what is easy to measure is not necessarily what is desirable to measure. It is always tempting to set goals based on the data that are gathered, rather than developing a data-gathering system linked to assessing progress towards meeting establishing goals" (Kyrillidou, 1998, p. 6).

The organizational issues centered on direction, leadership, communication, and support remained barriers. Covey expanded on these themes in her Digital Library Federation (DLF) study, *Usage and Usability Assessment: Library Practices and Concerns* (2002):

The results of the DLF study suggest that individually, libraries in many cases are collecting data without really having the will, organizational capacity, or interest to interpret and use the data effectively in library planning. . . . Comments from DLF respondents indicate that the internal organization of many libraries does not facilitate the gathering, analysis, management and strategic use of assessment data. The result is a kind of purposeless data collection that has little hope of serving as a foundation, for the development of guidelines, best practices, or benchmarks. The profession could benefit from case studies of those libraries that have conducted research efficiently and applied the results effectively. Understanding how these institutions created a program of assessment—how they integrated assessment into daily library operations, how they organized the effort, how they secured commitment of human and financial resources, and what human and financial resources they committed—would be helpful to the many libraries currently taking an ad hoc approach to assessment and struggling to organize their effort. (Covey, 2002, p. 58)

We live in a data-rich information environment that too "often far outpaces the ability to consistently, conscientiously, effectively and efficiently interpret the data and apply the conclusions and recommendations into various real-life decision-making situations" (Young & Peters, 2003, p. 1753). The following sections review data collection and analysis issues as well as providing examples of academic libraries that have successfully used data in planning and management.

DATA COLLECTION, ANALYSIS, AND REPORTING ISSUES

What Data Should Be Collected?

Libraries have a long history of collecting data on the size of collections, expenditures, staff, and other input elements. Outputs (for example, tallies

of customer use of resources) are another part of the picture. Using ARL statistics as a guidepost, the only indicators of customer activity currently collected and reported are circulation, reference transactions, interlibrary loan transactions, and delivery of bibliographic instruction. All these data elements provide valuable information, but they are no longer deemed sufficient. In using data for management purposes, a multifaceted approach (or a "toolkit") is advisable. As noted at a meeting of the ARL Committee on Statistics and Measurement, "different sets of data may be meaningful for different sets of libraries" (ARL, 2001, p. 3). It is only through using a wide variety of measures that one can hope to get a full and accurate reading of the library's activities.

Since the early 1990s a number of libraries have moved beyond mere counting of customer activity to carrying out surveys to learn about their customers. Are the customers satisfied with the delivery of service? Is the library offering the right services? What would customers like to see in the future? Hiller and Self describe the series of independently developed customer surveys conducted at their own institutions, the University of Washington and the University of Virginia (Hiller & Self, 2002).

In 1994 ARL adopted a new goal, to "describe and measure the performance of research libraries and their contributions to teaching, scholarship and community service" (Kyrellidou, 1998, p. 8). This was the start of the ARL New Measures Initiative, which formally began activity in 1999. This initiative would inform data collection that would go beyond traditional input/output measures to capture use and impact of libraries. In 1999 eight areas of interest were identified: user satisfaction, market penetration, ease and breadth of access, library impact on teaching and learning, library impact on research, cost effectiveness of library operations and services, library facilities and space, and organizational capacity (Blixrud, 2003).

In the past five years many libraries have chosen to participate in a specialized survey called LibQUAL+™. This survey is an adaptation of SERVQUAL™, a service quality assessment tool introduced in 1988 by the marketing team of Berry, Parasuraman, and Zeithaml (Nitecki, 1997; Zeithaml, Parasuraman, & Berry, 1990). During the 1990s the library at Texas A&M University pioneered the use of SERVQUAL™ in libraries. In 2000 Texas A&M and ARL began a joint project to adapt SERVQUAL for use in multiple libraries. Twelve ARL libraries participated in the development and testing of the instrument during the project's first year. The program, since named LibQUAL+™, has grown exponentially and now includes over 400 libraries of all sizes throughout the world (ARL 2003).

LibQUAL+™ is a gap analysis tool. Colleen Cook of Texas A&M explains: "It undertakes to measure library users' perceptions of service quality and identifies gaps between desired, perceived, and minimum expectations of service" (Cook, Heath, Thompson, and Thompson, 2001, p. 265). The instrument is designed to be useful to the library administration on several

levels: identifying deficits in service performance at an individual library, allowing comparisons with cohort libraries from multiple perspectives, identifying best practices, and responding to pressures for accountability. Cullen has voiced some reservations about LibQUAL+™; she acknowledges its value as a diagnostic tool and for longitudinal comparisons, but she questions its appropriateness for interinstitutional comparisons (Cullen, 2002).

Many libraries are also conducting studies of their internal processes and developing performance standards. These studies look for answers to questions such as the following: How fast is a service provided? What is the turnaround time for filling a user request? What is the error rate? The cost effectiveness of services and resources is also worthy of study. What is the cost per use of a given electronic journal? What does it cost the library to secure a book on interlibrary loan?

Usability testing is another area of inquiry. Jeffrey Rubin defines usability testing as "the process that employs participants who are representative of the target population to evaluate the degree to which a product meets specific usability criteria" (Rubin, 1994, p. 25). Libraries are now offering many, if not most, of their products and services in an online Web-based mechanism. In recent years libraries (and other service providers) have begun to realize that the design of Web sites can greatly affect their functionality. The best way to find out if a Web site is usable is to observe actual users as they attempt to use it. In recent years numerous articles on usability have appeared in the library press, and in 2001 the Library and Information Technology Association, a division of the American Library Association, published a collection of case studies related to usability (Campbell, 2001).

A number of libraries have made efforts to improve their services and processing by learning from peer institutions. They have conducted benchmarking or best practices projects, observing other institutions and changing their own practices as appropriate (Pritchard, 1995; White, 2002). St. Clair points out that a benchmarking project can improve the efficiency, effectiveness, and credibility of an organization, but it should not be undertaken lightly. "Benchmarking is a complex process requiring a genuine search for improvement on the part of the initiating institution. A significant investment of time must be made" (St. Clair, 1997, pp. 210–211).

Libraries are also moving beyond input and output measures by focusing on outcomes assessment. The purpose is to determine what impact the library has on the life of its clientele. In the felicitous words of Roswitha Poll, there is a need to measure the "preciousness of library services" (Poll, 2003). The College Libraries Section Standards Committee of the Association of College and Research Libraries (ACRL) offers a more prosaic definition: "Outcomes are the ways in which library users are changed as a result of their contact with the library's resources and programs" (ACRL, 2000).

Cullen notes that outcomes assessment is not a fully mature discipline: "Outcomes have proved to be a more difficult area of evaluation, and there is no work to date on standards for outcomes" (Cullen, 2002, p. 9).

In the United States, the Institute for Museum and Library Services (IMLS), a major source of federal funding for museums and libraries, has become an active proponent of outcomes assessment, educating practitioners through publications (Rudd, 2000) and presentations at national library conferences (IMLS/PLA, 2002). Further, the IMLS is now asking funding recipients to utilize outcomes-based evaluation. The IMLS Web site explains, "A focus on measuring outcomes—the effect of an institution's activities and services on the people it serves—rather than on the services themselves (outputs) is an emerging keystone of library and museum programs" (Sheppard, n.d.).

Prioritizing the Data

Research libraries are large organizations capable of generating an immense amount of complex data. The question inevitably arises as to the utility of the various data elements. *Which data should be collected and reported?* Libraries routinely collect the data requested by national organizations (for example, ARL, ACRL) and government agencies (for example, the Academic Libraries Survey of the National Center for Education Statistics), but these data elements may no longer be very useful to the individual libraries. Some of the organizations are revising their data collection priorities, hoping to increase the utility of the statistics. A number of individual libraries are engaged in a similar process, trying to determine what statistics they need.

One response to the flood of data is to identify the important data elements (those most crucial to the mission of the library) and to tally them as part of an overall index or scorecard. A few libraries have begun using an instrument called the Balanced Scorecard. This tool was developed in the United States in the early 1990s by two professors at the Harvard Business School; it was designed for the private sector, but more nonprofit and government agencies are now making use of it (Kaplan & Norton, 1992). The balanced scorecard allows a library to concentrate on a small number of measures. Taken together, these measures provide a quick but comprehensive picture of the health of the organization. The measures are divided into four categories, or perspectives: users, finance, internal processes, and learning and the future. Each perspective contains four to eight measures, and each measure includes a specific target score. At the end of the measurement period there should be no question as to which measures have met their targets.

Klaus Ceynowa from the University and Regional Library of Muenster (Germany) notes the strength of this tool: "The balanced scorecard compels the library management to concentrate on the evaluations critical to suc-

cess in the quality, cost efficiency and promptness of university information provision" (Ceynowa, 2000, p. 163). In North America the University of Virginia Library has been a notable proponent of the balanced scorecard; its activities have been chronicled at conferences and in publications (Oltmanns & Self, 2002; FLICC, 2003; Self, 2003a, 2003b).

Assessing Electronic Resources

In the past two decades libraries have undergone a virtual revolution in the variety of resources and services offered. Bibliographic databases have superseded print indexes in libraries. In academic libraries print journals are in some danger of being eclipsed by the electronic versions. An extensive review of electronic library use research has recently been published by Tenopir (2003).

King and colleagues (2003) present evidence of this movement from print to electronic journals. They conducted a set of readership surveys from 2000 to 2002 among four distinct groups of scientists and found that scientists with wide access to electronic journals tended to rely on and prefer the electronic format. Goodman (2002) notes that at Princeton users insisted that the library offer journals in electronic format. He also reports that introduction of journals in electronic format appears to result in a doubling of use. A large study at the University of California, as reported by Schottlaender, revealed that "digital use exceeded print use by at least an order of magnitude" (Schottlaender, 2003, slide 2 notes). He also expects an extended period of transition: "the care and feeding of hybrid collections of print and digital content is likely to be with us for some time to come" (Schottlaender, 2003, slide 17 notes). ARL statistics show a rapid increase in the proportion of the collections budget devoted to the purchase and licensing of electronic materials. The latest figures indicate that the typical ARL library in 2002 spent 21 percent of its collections budget on electronic materials.

The use of these digital materials can be tallied by the computer; in theory libraries should have more data, and more accurate data, than was ever possible with traditional materials. Usage data for electronic resources have enormous potential for assessment and management. Once libraries know which materials are being used, and how much each use costs, it becomes much easier to make selection decisions or to make a case for additional funding (Luther, 2002).

Unfortunately there are problems in utilizing the data. Libraries are dependent upon vendors for the data, and some vendors have declined to provide it. In addition there has been a lack of consensus as to what data elements should be counted. Moreover, even if vendors are counting the same things, they may not be counting them the same way (Blecic, Fiscella, & Wiberly, 2001; Davis, 2002). As noted in a recent article, "Perhaps one of the biggest reasons why it is difficult for libraries to use electronic re-

source data for decision making is the inconsistency across vendors" (Duy & Vaughan, 2003, p. 16). The difficulties in acquiring reliable and comparable data from different vendors, who may be in competition with one another, has led some libraries to develop their own methods for estimating use of electronic resources (Duy & Vaughan, 2003).

In the past few years there have been attempts to draft standards for usage statistics for electronic resources. Initiatives from both ARL and the International Coalition of Library Consortia (ICOLC) have resulted in suggestions or guidelines concerning electronic usage statistics. The latest and most promising effort is by an organization called COUNTER (Counting Online Usage of NeTworked Electronic Resources). The COUNTER Web site states the rationale for the organization:

The use of online information resources is growing exponentially. It is widely agreed by producers and purchasers of information that the use of these resources should be measured in a more consistent way. Librarians want to understand better how the information they buy from a variety of sources is being used; publishers want to know how the information products they disseminate are being accessed. An essential requirement to meet these objectives is an agreed international Code of Practice governing the recording and exchange of online usage data. COUNTER has developed just such a Code of Practice. (COUNTER, n.d., par. 1)

COUNTER is an international organization that counts among its membership the major national library organizations and a number of major publishers and aggregators. It has specified that statistical reports should contain certain data elements and that they should be presented in a specific, easy to use format. COUNTER also includes provisions for certification and auditing. As soon as a vendor is fully compliant with the guidelines, COUNTER certifies them and adds them to the official list of compliant vendors. This provision will definitely benefit libraries by clarifying standards for reporting. In summary, COUNTER is offering much hope that libraries will soon have data that are intelligible, reliable, and comparable. Such data may soon play a central role in the library assessment process.

The situation is less clear when it comes to locally owned or locally mounted digital materials. A number of libraries are engaged in extensive projects of building their own digital collections. Many of the major institutions in North America are listed on the Digital Library Federation Web site.¹ These libraries are digitizing varied materials and making them available to their clientele. Use of the materials may be completely unrestricted—anyone in the world may view and download them. In these cases, how should a library measure use of the material? There is no consistent practice. Libraries may count hits, page views, or sessions. All of these approaches have their adherents. There is also the question of *who* should be counted. Is remote use, perhaps from the other side of the world, counted

the same as use by an on-campus student? Are automated "visitors" (known as "spiders," "crawlers," or "bots") counted? Is inadvertent double clicking counted, or is it filtered out (Borghuis, 2000)? In the literature one can find a number of articles explaining how to extract data from Web log files, as well as other articles asserting that data taken from Web logs are useless (Bauer, 2000; Dowling, 2001; Goldberg, n.d.).

At the present time digital libraries seem to be in a developmental stage, with a focus on planning, creation, and experimentation. Only limited attention is given to assessment. However, Borgman and Larsen indicate the need for assessment: "A major challenge for digital library research and practice is to find relatively non-intrusive, low cost means of capturing appropriate data to assess the use, adoption, implementation, economics, and success of digital libraries" (Borgman & Larsen, 2003, par. 1). As digital libraries mature, assessment may well receive a higher priority.

The Limits of Quantitative Data

A word of caution is in order at this point. Although quantitative information is very powerful and important, it may not always be sufficient. There should be some feedback from users, some idea as to who is using a source. An undergraduate writing a paper can often substitute one source for another without harm; a research professor may not find it so easy to make a substitution if the first choice is not available. Therefore, librarians would be well advised to consider the opinions of faculty and other stakeholders, along with the quantitative data, as they make decisions about selection and retention of electronic resources. Qualitative information from focus groups, usability, and observation has been of immense use in understanding user behavior. Focus groups, in particular, have provided powerful context to enrich and complement other assessment efforts. The small group process encourages people to express their views, priorities, and concerns directly in a way that other methods cannot.

Gorman points out the limits and dangers of data collection; he encourages managers to supplement data with rich qualitative information. He is especially concerned that stakeholders, such as political and financial agencies, may misinterpret and misuse quantitative data. He argues for "a greater awareness among library professionals that meaningful data are contextual; and that meaning depends on interpretation" (Gorman, 2000, p. 118).

NOTABLE LIBRARIES

Many libraries deserve notice for their work with data and assessment. The following includes brief reports of two noteworthy libraries and first-hand reports from the authors' own institutions. Assessment activities at these four libraries have been widely reported at meetings of ARL, the American Library Association (ALA), ACRL, and other library organiza-

tions. As an example, representatives from each of these libraries will participate in a program at the 2004 ALA annual conference called "Best Practices: Collection Management and the Application of New Measures for Library Assessment" (ALCTS, 2003).

University of Arizona: Measuring Organizational Performance

In 1991 the University of Arizona responded to a fiscally challenging environment by hiring Carla Stoffle as library director. Since then the University of Arizona Library has been a leader in organizational innovation and the use of quantitative information to improve performance. Simultaneously Stoffle also served several years as chair of the Statistics and Measurement Committee of ARL; she has been influential in moving both organizations into new areas of assessment and measurement.

The Arizona experience has been widely reported.² The University Library also cosponsors a biennial conference called "Living the Future," which features innovations at Arizona and other libraries: "We wanted to share with colleagues our own successes and challenges as we transformed from an academic library of the 20th century into one that is preparing for the next millennium."³

In an overview of their work at Arizona, Stoffle and Phipps note the importance of implementing an organizational performance measurement system. In 1998 the Library formally adopted a system known as PEMS (Performance Effectiveness Management System). PEMS is based on continual assessment of client needs, and it includes standards or targets for each activity. It is part of a cultural change that has taken place: "Developing a system approach to measurement helps develop an internal culture of assessment where decisions are guided by facts, research, and analysis" (Stoffle & Phipps, 2003, p. 26). Veldof described a number of assessment and data collection methods used to improve performance and summarized:

For the University of Arizona, data-driven research did indeed matter and continues to matter on a daily basis. Data and its collection and analysis are catalysts to move libraries to make changes, to measure their progress towards these changes, to direct efforts in ways that will give libraries the biggest impact for the lowest cost, and ultimately to greatly improve customer satisfaction. (Veldof, 1999, p. 32)

University of Pennsylvania: Dynamic Data

The University of Pennsylvania Library, another leader in the collection and presentation of quantitative data, utilizes an extremely interactive approach. Mounted on the library's Web site is a repository of quantitative information called the Penn Library Data Farm. Its stated purpose is "to aid the measurement and assessment of library resource use and organizational performance" (University of Pennsylvania Library, n.d., par. 1). The Data Farm includes both locally produced data and vendor statistics

and the software needed to produce reports. It is described as "not a static warehouse of figures, but a more dynamic program . . . that equips staff to analyze and assess their work independently" (University of Pennsylvania Library, n.d., par. 1).

The Data Farm is an effort to provide "one-stop shopping" for library data. The site includes circulation and gate counts, database and e-journal use, survey reports, and various other data elements. It also allows one to run a variety of customized programs that process transaction logs, update use statistics, and generate Web use reports (Zucca, 2003b).

In tallying the use of electronic resources, Penn has chosen not to rely on statistics provided by the vendors. Zucca points out many of the problems with vendor statistics, for example, useless or low-resolution data, no consensus of definitions, no uniformity of metrics, and difficulty in retrieving data. He then describes Penn's strategy for overcoming the problems: "Gather consistent, clearly defined measures for all e-resource use, based on mechanisms available to and controlled by Penn Library" (Zucca, 2003a, slide 7). To accomplish this strategy Penn has built measurement devices into its local architecture and created tools for storing, organizing, normalizing, and processing the collected data. Zucca (2003a) notes that much of the impetus for the Data Farm was external, especially as it relates to the library's ability to justify use as a cost center. He provides the following reasons for development of data-based assessment at the Penn Library:

- Responsibility center budgeting: tax the schools for central services
- Expectation of flat or declining budgets
- High standards of accountability for cost centers
- Provost is a quantitative scientist
- Empirical mindset of library leadership

University of Virginia: Using Data to Inform Decision-Making

The University of Virginia Library has a long history of utilizing statistics and quantitative data analysis as part of its effort to provide high-quality services. Kendon Stubbs, recently retired as deputy university librarian, was a leader in the development of ARL statistical initiatives (Stubbs, 1981). In the 1980s the Library conducted a large-scale investigation of the effect of reserve use on student grades (Self, 1987). In the 1990s a two-year study of circulation patterns of newly acquired monographs led to a drastic change in the collection development policies and a reorganization of collection development functions within the Library (Self, 1996).

The Library administration formalized its commitment to data collection and analysis in 1991 when it established the Management Information Systems Committee. Among other tasks, the committee was asked to serve as a clearinghouse for computer programs that generate management data, to identify areas within the library where performance could be enhanced with management data, and to educate staff in topics relating to manage-

ment information. In 1993 the committee moved into a new area with the implementation of a comprehensive survey of teaching and research faculty. The following year the committee carried out a similar survey of undergraduates and graduate students. Surveys have continued on an approximately triennial basis, with faculty surveys in 1996, 2000, and 2003, and student surveys in 1998 and 2001. The first two surveys were on paper, but since 1996 the surveys have been administered and published on the World Wide Web.

From the outset the Library worked to maximize the response rates for the surveys. Faculty response rates have ranged from a low of 62 percent to a high of 70 percent; among graduate students the rates have been between 50 percent and 60 percent, and among undergraduates from 40 percent to 50 percent. The relatively high response rates have enabled the Library to use the results with some assurance of their reliability. The administration of the Library has been able to use the survey results to support implementation of innovative services, for example, an electronic text center, a Web-based library catalog, a coffee shop in the library, and a transition toward electronic acquisitions.

Survey results are particularly useful when they can be corroborated with other data. One example at Virginia concerned activity at the traditional reference desk. Tallies of reference desk queries had been in decline for several years, but there was controversy as to whether the tallies were reliable or merely a statistical fluke. However, the tallies correlated closely with a longitudinal analysis of answers on undergraduate surveys. It became clear that fewer undergraduates were making use of traditional reference services, and the Library was able to adjust staffing patterns accordingly.

In 2000 the practice of management information services moved to a new level at Virginia. The committee was disbanded and replaced by a three-person MIS department. The new department has responsibility for assessment, data collection and reporting, and usability testing, as well as certain budgetary tasks. Volunteers from various departments continue to participate in these activities, but coordination is the responsibility of the MIS department.

University of Washington: User Needs Assessment

The University of Washington Libraries (UW Libraries) is known for its extensive work in user needs assessment (Hiller 2001, 2002b, 2002c). Since the first large-scale faculty and student surveys in 1992, the UW Libraries, with strong administrative support and broad-based staff participation, has conducted extensive, ongoing assessment work with the user community, focusing on user needs assessment, priorities, library and information use patterns, and user satisfaction with the quality of library services and collections. The UW Libraries has employed a variety of methods to obtain information from faculty and students, including large-scale surveys, tar-

geted surveys, focus groups, observation studies, usability testing, guided interviews, meetings, and both traditional and electronic suggestion boxes. Assessment results guide and inform the development and improvement of services and resources that support the academic community.

The library assessment coordinator (half-time) chairs the Library Assessment Committee, and together they coordinate broad-based assessment efforts and provide knowledge support for other library-related assessment activities. All areas within the Libraries are encouraged and supported to incorporate assessment and evaluation into their ongoing activities and programs, so that assessment becomes a routine part of library operations. Indeed, the phrase "culture of assessment," which is widely used within the library community to define an institution where assessment is an ongoing, ingrained activity, was first coined at the UW Libraries in 1994.

The UW Libraries program of user surveys is unique among academic research libraries. Since 1992 large-scale surveys of students and faculty are conducted on a three-year cycle. These triennial surveys have provided invaluable information about how students and faculty use libraries, their library and information needs and priorities, and the importance of and satisfaction with the Libraries during a period of rapid change in the information environment. The large number of faculty respondents (1300–1500 per survey) is sufficient to conduct analysis below the aggregate level at the school and college level. Survey instruments, results, and basic analysis are provided on the Libraries Assessment Web site (University of Washington Libraries, n.d.)

The UW Libraries has used survey results to improve library services and programs based on customer needs. These changes have included renovating library facilities for student use and reducing collections space; extending hours for branch libraries and providing twenty-four-hour access for the undergraduate library; installing more library computers for student use; moving rapidly to desktop delivery of full-text resources; identifying student information technology support needs and working with campus partners to address them; providing standardized service training for library staff and student assistants who work directly with the user community; consolidating and merging branch libraries; understanding that information and library needs differ between groups and academic areas and planning services tailored to these needs.

Other broad-based surveys include participation each year in the Association of Research Libraries sponsored LibQUAL+™ surveys and a cycle of in-library use surveys conducted on a triennial basis since 1993 to determine which groups use the physical library and why they visit. LibQUAL+™ is a cost-effective complement to the library's own surveys and results can also be compared with peer institutions. Use of in-library survey data has ranged from developing a service levels policy in 1995 that articulated services to

nonaffiliated users, to expanding the functionality of library computers in 2003 by adding application software.

Since 1998 the UW Libraries has conducted focus groups on an annual basis with faculty and students. Focus group discussions have identified potential problem areas and have led to action. For example, 2003 focus groups on the topic of information literacy confirmed the complexity of the information environment and the difficulties students and faculty have in finding scholarly information for their work. These findings helped initiate a redesign process for the Libraries Web site that will facilitate resource discovery and also intensify library efforts to integrate information literacy into curricular design.

The UW Libraries assessment efforts and ability to use results to improve library services were recognized in the decennial accreditation review of the university in 2003:

In view of the overall excellence of the Libraries, it should not be surprising that they have benefited from having visionary leaders. Planning, assessment, and continuous improvement are ongoing processes with broad staff participation. The Libraries' program for the measurement of library use and user satisfaction has resulted in 10 years of longitudinal data on satisfaction rates and user behavior. This information is frequently referred to and used to modify existing services and plan new ones. (Northwest Association of Schools and of Colleges and Universities, 2003, III-5-1)

CONCLUSION

The outlook for the effective use of data in library planning and management is far more optimistic now than five or ten years ago. Not only are successful programs in place at several libraries that can serve as realistic models, but the emergence of a robust support infrastructure provides the guidance and expertise that can help develop and sustain data-based decision-making. Blixrud notes that, for institutions to do measurement and evaluation effectively, it takes

- Resources (that is, time and money)
- Individual and institutional buy-in
- Access to individuals to evaluate
- Expertise to conduct evaluation
- Project management experience
- Appropriate benchmarks
- Conceptual clarity
- Measurement and design requirements
- Instrument validity and reliability (Blixrud, 2003, p. 7)

Fortunately, many of these points are now being addressed. ARL has taken the lead in developing or sponsoring programs that assist libraries

(not just research libraries) in developing the skill base and organizational climate for effective support. Such programs as the Service Quality Evaluation Academy, a week-long workshop on how to use and understand quantitative and qualitative data, or the online lyceum on measuring library service quality, have reached many. The widespread use of LibQUAL+™ across a broad spectrum of more than 400 libraries has done much to foster a “culture of assessment” and the collection of relevant user data. Other relevant ARL initiatives that are data-based include internal processes such as interlibrary loan and measuring use of electronic resources (Blixrud, 2003). ARL also contributes substantial programming for organizational and leadership development.

Other professional organizations have regularly included sessions and workshops on library evaluation, assessment, and data use as part of their conferences. The Northumbria International Conference on Performance Measurement in Libraries and Information Services has held five successful conferences since 1995, bringing together hundreds of library educators, researchers, and practitioners to discuss how to measure quality and apply it in libraries. Library and information schools are equipping not only students with these skills but also expanding their continuing education efforts for practitioners.

External factors such as accreditation, accountability, and financial support play ever larger roles in demanding that libraries demonstrate they operate effectively and efficiently in addressing the needs of their communities. Gratch-Lindauer (2002) notes that many accrediting agencies have moved from using input measures for libraries to outcomes-based assessment. Libraries that have established an integrated program of acquiring and using data wisely will be better positioned to meet these challenges.

Organizational development provides the structure for libraries to “institutionalize” the routine collection and use of data in planning and management. Administrative leadership and support are critical to fostering an environment that equips staff with the vision, tools, and understanding necessary to make data-based decision-making an integral organizational value. Yet, there is no one way to attain this. Each of the four libraries discussed above takes a different approach to achieving this value, one that is aligned with the culture, mission, goals, and objectives of each institution. It is also important to recognize that each of these libraries did not wait until all the organizational pieces were in place but started incrementally. In the long run, success will be measured by how effective each library is in using data wisely to provide the services and resources that support the learning, research, and information needs of its community.

NOTES

1. See <http://www.diglib.org/>.
2. See "UA Library Organization Articles" at <http://www.library.arizona.edu/library/teams/fast/biblio.html>.
3. See <http://www.library.arizona.edu/conference/about.html>.

REFERENCES

- Allen G. (1985). The management use of library statistics. *IFLA Journal*, 11(3), 211–222.
- Association of College & Research Libraries (ACRL). (2000). *Standards for college libraries 2000 edition*. Retrieved January 2, 2004, from http://www.ala.org/Content/NavigationMenu/ACRL/Standards_and_Guidelines/Standards_for_College_Libraries_2000_Edition.htm.
- Association for Library Collections & Technical Services (ALCTS). (2003). *Best practices: Collection management and the application of the new measures for library assessment*. Retrieved January 5, 2004, from http://www.ala.org/Content/NavigationMenu/ALCTS/Continuing_Education2/Events11/ALCTS_Annual_Conference_Programs/Collection_Management_and_the_Application_of_New_Measures_for_Library_Assessment.htm.
- Association of Research Libraries (ARL). (1987). *Planning for management statistics in ARL libraries* (SPEC Kit No.134). Washington, DC: Office of Management Studies, Association of Research Libraries.
- Association of Research Libraries (ARL). (1989a). *Use of management statistics in ARL libraries* (SPEC Kit 153). Washington, DC: Office of Management Studies, Association of Research Libraries.
- Association of Research Libraries (ARL). (1989b). *Strategic plans in ARL libraries* (SPEC Kit No. 158). Washington, DC: Office of Management Studies, Association of Research Libraries.
- Association of Research Libraries (ARL). (2001, May 23). *138th ARL Membership Meeting, ARL Committee on Statistics and Measurement* [meeting minutes]. Toronto, Canada. Retrieved January 2, 2004, from <http://www.arl.org/stats/program/2001/a2-min1005.pdf>.
- Association of Research Libraries (ARL). (2003). LibQUAL+ History. Retrieved January 2, 2004, from <http://www.libqual.org/About/History/index.cfm>.
- Bauer, K. (2000). Who goes there? Measuring library Web site usage. *Online*, 24(1), 25–26, 28, 30–31.
- Blecic, D. D., Fiscella, J. B., & Wiberly, S. E. (2001). The measurement of use of Web-based information resources: An early look at vendor-supplied data. *College & Research Libraries*, 62(5), 434–453.
- Blixrud, J. (2003). Mainstreaming new measures. *ARL: A Bimonthly Report on Research Library Issues and Actions*, 230/231, 1–8.
- Borghuis, M. (2000, September 20). What to count and what not? A white paper on the filters to be applied to a Web-server log file before usage-analysis and reporting can start (Final Version). Amsterdam: Elsevier ScienceDirect.
- Borgman, C., & Larsen, R. (2003, September). In brief: ECDL 2003 Workshop report: Digital library evaluation—Metrics, testbeds and processes. *D-Lib Magazine*, 9. Retrieved January 5, 2004, from <http://www.dlib.org/dlib/september03/09inbrief.html#BORGMAN>.
- Campbell, N. (2001). *Usability assessment of library-related Web sites: Methods and case studies*. Chicago: Library Information Technology Association.
- Ceynowa, K. (2000). Managing academic information provision with the Balanced Scorecard: A project of the German Research Association. *Performance Measurement and Metrics*, 1(3), 157–164.
- Clapp, V., & Jordan R. (1965). Quantitative criteria for adequacy of academic library collections. *College and Research Libraries*, 26, 371–380.
- Cook, C., Heath, F., Thompson, B., & Thompson, R. L. (2001). LibQUAL+™: Service quality assessment in research libraries. *IFLA Journal*, 27(4), 264–268.
- COUNTER. (n.d.). *Introducing COUNTER*. Retrieved January 2, 2004, from <http://www.projectcounter.org/index.html>.
- Covey, D. (2002). *Usage and usability assessment: Library practices and concerns*. Washington, DC: Digital Library Federation, Council on Library and Information Resources.

- Cullen, R. (2002). Setting standards for service quality. In J. Stein, M. Kyrrilidou, & D. Davis (Eds.), *Meaningful measures for emerging realities: Proceedings of the Fourth Northumbria International Conference on Performance Measurement in Libraries and Information Services* (pp. 9–16). Washington, DC: Association of Research Libraries.
- Davis, P. M. (2002). Patterns in electronic journal usage: Challenging the composition of geographic consortia. *College & Research Libraries*, 63(6), 484–497.
- DeProspero, E., Altman, E. R., & Beasley, K. E. (1973). *Performance measures for public libraries*. Chicago: Public Library Association.
- Dowlin, K., & McGrath, L. (1983). Beyond the numbers—A decision support system. In F. W. Lancaster (Ed.), *Library automation as a source of management information/Clinic on library applications of data processing* (pp. 27–58). Champaign, IL: Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign.
- Dowling, T. (2001). Lies, damned lies, and Web logs. *School Library Journal*, 47(5), 34–35.
- Duy, J., & Vaughan, L. (2003). Usage data for electronic resources: A comparison between locally collected and vendor-provided statistics. *Journal of Academic Librarianship*, 29(1), 16–22.
- FLICC. (2003). FLICC/FEDLINK program evaluation: The Balanced Scorecard method. [video presentation, segment II, February 14, 2003]. Washington, DC: Library of Congress. Retrieved January 2, 2004, from <http://www.loc.gov/flicc/video/balance/balancedscore.html>.
- French, W., & Bell C. (1999). *Organization development: Behavioral science interventions for organization improvement* (6th ed.). Upper Saddle River, NJ: Prentice Hall.
- Gerould, J. T. (1906). A plan for the compilation of comparative university and college library statistics. *Library Journal*, 31, 761–763.
- Goldberg, J. (n.d.). *Why Web usage statistics are worse than meaningless*. Retrieved January 2, 2004, from <http://www.goldmark.org/netrants/webstats/>.
- Goodman, D. (2002). A year without print at Princeton, and what we plan next. *Learned Publishing*, 15, 43–50.
- Gorman, G. E. (2000). Collecting data sensibly in information settings. *IFLA Journal*, 26(2), 115–119.
- Gratch-Lindauer, B. (2002). Comparing the regional accreditation standards: Outcomes assessment and other trends. *Journal of Academic Librarianship*, 28(1/2), 14–25.
- Heim, K. (1983). Organization considerations relating to the implementation and use of Management Information Systems. In F. W. Lancaster (Ed.), *Library automation as a source of management information/Clinic on library applications of data processing* (pp. 59–70). Champaign, IL: Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign.
- Hernon, P., & McClure, C. (1990). *Evaluation and library decision making*. Norwood, NJ: Ablex Publishing.
- Hiller, S. (2001). Assessing user needs, satisfaction and library performance at the University of Washington Libraries. *Library Trends*, 49(4), 605–625.
- Hiller, S. (2002a). “But what does it mean?” Using statistical data for decision making in academic libraries. In C. Creaser (Ed.), *Statistics in practice—Measuring and managing: Proceedings of IFLA Satellite Conference* (pp. 10–23). Loughborough, England: Loughborough University.
- Hiller, S. (2002b). The impact of information technology and online library resources on research, teaching and library use at the University of Washington. *Performance Measurement and Metrics*, 3(3), 134–139.
- Hiller, S. (2002c). How different are they? A comparison by academic area of library use, priorities, and information needs at the University of Washington. *Issues in Science and Technology Librarianship*, 33. Retrieved January 2, 2004, from <http://www.isrl.org/02-winter/article1.html>.
- Hiller, S., & Self, J. (2002). A decade of user surveys: Utilizing and assessing a standard assessment tool to measure library performance at the University of Virginia and University of Washington. In J. Stein, M. Kyrrilidou, & D. Davis (Eds.), *Proceedings of the Fourth Northumbria International Conference on Performance Measurement in Libraries and Information Services: Meaningful measures for emerging realities* (pp. 253–261). Washington, DC: Association of Research Libraries.

- Institute of Museum and Library Studies/Public Library Association (IMLS/PLA). (2002). *Outcomes: Libraries change lives—Oh yeah? Prove it*. Public Library Association, Phoenix, AZ. Retrieved January 5, 2004, from <http://www.ims.gov/grants/current/pla-02-2obe.pps>.
- Kaplan, R. S., & Norton, D. P. (1992). The Balanced Scorecard—Measures that drive performance. *Harvard Business Review*, 70(1), 71–79.
- King, D. W., Boyce, P. B., Montgomery, C. H., & Tenopir, C. (2003). Library economic metrics: Examples of the comparison of electronic and print journal collections and collection services. *Library Trends*, 51(3), 376–400.
- Kyrillidou, M. (1998). An overview of performance measures in higher education and libraries. *ARL: A Bimonthly Newsletter of Research Library Issues and Actions*, 197, 3–7.
- Lakos, A. (1999). The missing ingredient—Culture of assessment in libraries: Opinion piece. *Performance Measurement and Metrics*. Sample Issue August 1999. Retrieved January 5, 2004, from <http://www.aslib.com/pmm/1999/aug/opinion.html>.
- Lakos, A. (2002). Culture of assessment as a catalyst for organizational culture change in libraries. In J. Stein, M. Kyrillidou, & D. Davis (Eds.), *Proceedings of the Fourth Northumbria International Conference on Performance Measurement in Libraries and Information Services: Meaningful measures for emerging realities* (pp. 311–319). Washington, DC: Association of Research Libraries.
- Lakos, A., Phipps, S., & Wilson, B. (2002). *Defining a "culture of assessment"*. Retrieved January 13, 2004, from <http://www.library.ucla.edu/yr1/reference/aalakos/assessment/CulAssessToolkit/Assessdef3-new.doc>.
- Lancaster, F. W. (1977). *The measurement and evaluation of library services*. Washington, DC: Information Resources Press.
- Lancaster, F. W., & McCutcheon, D. (1978). Some achievements and limitations of quantitative procedures applied to the evaluation of library services. In C. Chen (Ed.), *Quantitative measurement and dynamic library service* (pp. 12–30). Phoenix, AZ: Oryx Press.
- Luther, J. (2002). Getting statistics we can use. In J. Stein, M. Kyrillidou, & D. Davis (Eds.), *Proceedings of the Fourth Northumbria International Conference on Performance Measurement in Libraries and Information Services: Meaningful measures for emerging realities* (p. 321). Washington, DC: Association of Research Libraries.
- McClure, C. R. (1984). Management information for library decision-making. *Advances in Librarianship*, 13, 1–47.
- McClure, C. R. (1986a). Preparing the organization for microcomputer-based decision making. In P. Hernon & C. McClure (Eds.), *Microcomputers for library decision making*. Norwood, NJ: Ablex, 39–57.
- McClure, C. R. (1986b, July). A view from the trenches: Costing and performance measures for academic library public services. *College & Research Libraries*, 47, 323–336.
- Metcalfe, K. (1986). *Planning academic and research library buildings* (2nd ed. by P. Leighton & D. Weber). Chicago: American Library Association.
- Molyneux, R. E. (1986). *The Gerould statistics*. Washington, DC: Association of Research Libraries.
- Nitecki, D. A. (1997). SERVQUAL: Measuring service quality in academic libraries. *ARL Bimonthly Report*, 191, 14.
- Northwest Association of Schools and of Colleges and Universities. (2003). *Evaluation Committee Report. University of Washington, Commission on Colleges*. Retrieved October 26, 2003, from <http://www.washington.edu/about/accreditation/EvaluationCommReport.pdf>.
- Oltmanns, G., & Self, J. (2002). *Assessment and flexibility: Implementing the Balanced Scorecard, reassigning staff*. Presentation at the Living the Future 4 Conference, Tucson, April 26. Retrieved January 2, 2004, from <http://www.library.arizona.edu/conference/ltf4/pres/welcome.html>.
- Poll, R. (2003). *Measuring impact and outcome*. Muenster, Germany: Regional and University Library. Retrieved January 2, 2004, from <http://www.uni-muenster.de/ULB/bibliothek/projekte/outcome.html>.
- Pritchard, S. (1995). Library benchmarking: Old wine in new bottles? *Journal of Academic Librarianship*, 21, 491–495.
- Riggs, D. E. (1984). *Strategic planning for library managers*. Phoenix, AZ: Oryx Press.

- Rubin, J. (1994). *Handbook of usability testing: How to plan, design, and conduct effective tests*. New York: John Wiley & Sons.
- Rudd, P. (2000). Documenting the difference: Demonstrating the value of libraries through outcome measurement. In *Perspectives on outcome based evaluation for libraries and museums* (pp. 16–23). Washington, DC: Institute of Museum and Library Services. Retrieved January 5, 2004, from <http://www.imls.gov/pubs/pdfs/pubobe.pdf>.
- Schottlaender, B. E. C. (2003). *University of California Collection Management Initiative: A study of behavior and attitudes of people using electronic journals, 142nd annual meeting, May 2003*. Lexington, KY: Association of Research Libraries. Retrieved January 2, 2004, from http://www.ucop.edu/cmi/pubs_arl.html.
- Self, J. (1987, January). Reserve readings and student grades: Analysis of a case study. *Library & Information Science Research*, 9, 29–40.
- Self, J. (1996). 1993 baseline study: University of Virginia Library. Retrieved January 20, 2004, from <http://www.lib.virginia.edu/mis/reports/BaselineREV.pdf>.
- Self, J. (2003a). From values to metrics: Implementation of the Balanced Scorecard at a university library. *Performance Measurement and Metrics*, 4(2), 57–63.
- Self, J. (2003b). Using data to make choices: The Balanced Scorecard at the University of Virginia Library. *ARL: A Bimonthly Report on Research Library Issues and Actions*, 230/231, 28–29.
- Shank, R. (1983). Management, information and the organization: Homily from the experience of the data rich but information poor. In F. W. Lancaster (Ed.), *Library automation as a source of management information/Clinic on library applications of data processing* (pp. 2–9). Champaign, IL: Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign.
- Sheppard, B. (n.d). *Showing the difference we make: Outcome evaluation in libraries and museums*. Institute of Museum and Library Services. Retrieved January 2, 2004, from http://www.imls.gov/grants/current/crnt_obe.htm#intro.
- St. Clair, G. (1997). Benchmarking and restructuring at Penn State Libraries. In C. A. Schwartz (Ed.), *Restructuring academic libraries* (pp. 200–212). Chicago: Association of College & Research Libraries.
- Stoffle, C. J., & Phipps, S. (2003). Creating a culture of assessment: The University of Arizona experience. *ARL Bimonthly Report*, 230/231, 26–27.
- Stoffle, C. J., Renaud, R., & Veldof, J. R. (1996). Choosing our futures. *College & Research Libraries*, 57, 213–33.
- Stubbs, K. (1981). University libraries: Standards and statistics. *College & Research Libraries*, 42(6), 528–38.
- Swisher, R., & McClure, C. R. (1984). *Research for decision making: Methods for librarians*. Chicago: American Library Association.
- Tenopir, C. (2003). *Use and users of electronic library resources: An overview and analysis of recent research studies*. Washington, DC: Council of Library and Information Resources. Retrieved, January 17, 2004, from <http://www.clir.org/pubs/reports/pub120/pub120.pdf>.
- University of Pennsylvania Library. (n.d.). *Penn Library Data Farm*. Retrieved January 2, 2004, from <http://metrics.library.upenn.edu/prototype/about/index.html>.
- University of Washington Libraries. (n.d.). *UW Libraries Assessment*. Retrieved January 2, 2004, from <http://www.lib.washington.edu/assessment/>.
- Van House, N. A., Weil, B., & McClure C. R. (1990). *Measuring academic library performance: A practical approach*. Chicago: American Library Association.
- Veldof, J. (1999). Data driven decisions: Using data to inform process change in libraries. *Library & Information Science Research*, 21(1), 31–46.
- White, L. (2002). The University of Virginia's experiment with benchmarking. *Virginia Libraries*, 48(4), 17–25.
- Young, A., & Peters, T. (2003). Management information systems in libraries. In M. A. Drake (Ed.), *Encyclopedia of library and information science* (pp. 1748–1756). New York: Marcel Dekker.
- Zeithaml, V. A., Parasuraman, A., & Berry, L. L. (1990). *Delivering quality service: Balancing customer perceptions and expectations*. New York: Free Press.
- Zucca, J. (2003a). *Find a way or make one: Strategies for measuring and analyzing the use of digital*

information. [ARL Survey Committee/E-Metrics meeting, May 23, 2001, Toronto, Canada]. Retrieved January 2, 2004, from <http://digital.library.upenn.edu/pubs/>.

Zucca, J. (2003b). Traces in the clickstream: Early work on a management information repository at the University of Pennsylvania. *Information Technology and Libraries*, 22(4), 175-179.