ARL

SPECIAL DOUBLE ISSUE ON NEW MEASURES
TABLE OF CONTENTS

MAINSTREAMING NEW MEASURES

by Julia C. Blixrud, Assistant Executive Director, External Relations, ARL

Within the past four years, "new measures" has entered the lexicon of library performance and assessment literature. In the 1990s, escalating economic pressures and demands for accountability from governing boards, university administrators, and government agencies strongly encouraged ARL member libraries and many other libraries to find new ways to illustrate the contribution their collections and services provided to their constituencies. Traditional, input measures (i.e., size of collections, budget, and staff) were not meeting institutional accountability needs. Although efforts were in place to measure some forms of output—e.g., activities such as circulation, interlibrary borrowing, and group presentations— these, too, did not always provide a good or complete indication of library performance and the cost effectiveness of their services. In the 1990s, the ARL Statistics and Measurement Committee began discussing the need to move beyond the traditional measures, building on earlier work in organizational development conducted by ARL's Office of Management Services in the late 1980s and throughout the 1990s.[1] In early 1999, the ARL New Measures Initiative officially began to develop measures that would better demonstrate libraries' institutional value and describe their current operations.

Now after four years of intensive work, the New Measures Initiative is no longer new. ARL member libraries have integrated some of the measures into the mainstream of library operations and the Statistics and Measurement program has an active agenda for research, experimentation, and demonstration projects.

Development of the Initiative

ARL has been collecting descriptive data from its members for dozens of years and a robust data file covers 1907–08 to the present. However valuable these data are for describing long-established characteristics of research libraries and for monitoring trends, they are not adequate to address emerging uses of technology for access to information nor are they helpful for describing the changing nature of service in research libraries. Over the years ARL experimented with gathering data to assess library performance. As early as 1982, ARL began a Supplementary Statistics form to test the gathering of different statistics, and if the variables proved useful, they were moved into the main Statistics. For example, collecting data on government documents and reference transactions was tested in the Supplementary Statistics and those items were added to the main Statistics in 1990–91 and 1994–95 respectively.

It was through the Supplementary Statistics test bed that ARL began collecting information on the character of research library investment in electronic resources. The definitional problems immediately became apparent, as did the difficulties for libraries in readily collecting some of the data. Between 1996 and 1998, Timothy Jewell (University of Washington) served as an ARL Visiting Program Officer to assess early efforts to collect measures of electronic resources. The goal of his inquiry was to develop standard definitions for more systematic collection of information about the transformation of research libraries. Through a series of meetings, discussions, reports, and analysis of existing data, the Jewell study documented a community seeking to measure facets of a complex environment subject to radical

shifts due to rapidly changing technology and applications. Not surprisingly, his final report highlights the wide range of efforts underway to capture this information but also two concerns: that any information gathered have clear utility, and that time and effort required to supply the information be minimized.

The tenor of library measurement discussions in the late 1990s was one of frustration peppered with examples of false starts in an environment that changed faster than it could be measured. There was a sense of urgency to develop better measures for the networked environment but there was no sense of commonality regarding which areas were most important to work on. What emerged from the experience of the late 1990s was a realization that any definition of new measures would need to be accepted by the community as part of the larger context of institutional assessment. There was also a realization that developing new measures carried costs and that this work had to be shared in order to minimize the expenses for any single institution.[2]

What emerged from the experience of the late 1990s was a realization that any definition of new measures would need to be accepted by the community as part of the larger context of institutional assessment.

ARL Director Retreat
Seeing the need for a common direction, Carla Stoffle (University of Arizona) and Paul Kobulnicky (University of Connecticut), chairs of the ARL Statistics and Measurement Committee and the Research Library Leadership and Management Committee, called a retreat in early 1999 in Tucson, Arizona. The goal of the retreat was to discuss how to develop strategies that would address the development of new measures.[3]

Retreat participants identified eight areas of interest for which measures would be particularly helpful:

- 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
- Organizational Capacity

In order for any specific new measures projects to succeed, it was determined that there must be collaboration among member leaders with strong interest in them; individual projects could and should be developed with different models for exploration; and there must be an intent to make resulting tools and methodologies available to the full ARL membership as well as the broader library community. Retreat participants agreed that not all ARL members would be required to participate in all projects, which would give each project the flexibility it needed to test and refine measures without placing undue burdens on the entire ARL membership. In addition, it was proposed that these new projects would not be used to modify the existing data-collection efforts without considerable testing and that the traditional measures would be "frozen" in their current form in order to retain their integrity. As new measures became available, the ARL Statistics and Measurement Committee would review each one for addition to the regular ARL data-collection activities.

After the retreat, brief working papers were developed by directors for some of the eight areas to expand on the ideas expressed and to identify possible approaches for investigation.[4] Authors were encouraged to build upon previous research to better ensure early success for the New Measures Initiative. The areas of interest were tested with the broader ARL community at the May 1999 Membership Meeting and plans were set in motion to begin developing the measures.

Progress in Areas of Interest
The Statistics and Measurement Committee, under the leadership of Carla Stoffle, directed the New Measures Initiative. An ambitious agenda was set forth directing attention to a number of areas. A series of projects and special events were initiated in 2000 and considerable progress has been made in nearly all of the areas of interest identified at the retreat.

User Satisfaction and Market Penetration
Combining two of the retreat's areas of investigation—reaching the library community and identifying the success libraries have serving users—resulted in the largest new measures activity: LibQUAL+™.[5] This research and development project undertaken by ARL in collaboration with Texas A&M University is defining and measuring library service quality across institutions and creating useful quality assessment tools. The Web-based, total-market survey measures library user perceptions of service quality and identifies gaps between minimum, desired, and perceived expectations of service. It provides project participants with reliable, tested, comparable data for internal and external benchmarking. In addition, a number of training activities have been developed to assist libraries in deploying the survey, interpreting results, and

implementing changes in their operations in response to the survey data. Colleen Cook and Fred Heath have led this effort from Texas A&M and Martha Kyrillidou and Duane Webster have led this effort from ARL. Two leading researchers from Texas A&M, Yvonna Lincoln and Bruce Thompson, have served as qualitative and quantitative experts respectively and the project's development has been well documented in the literature.[6]

Two events jumpstarted the LibQUAL+™ project. First, ARL held a symposium on measuring service quality in October 2000.[7] Experts in the field of library service quality and ARL members participating in a pilot project to test a library user survey based on SERVQUAL—a service quality evaluation tool used in the private sector—spoke about the need for an instrument that looked at library service from the user's point of view. The symposium raised interest in the LibQUAL+™ project specifically and the issue of measuring library service quality generally. Also in the fall of 2000, ARL and Texas A&M received a three-year grant from the U.S. Department of Education Fund for the Improvement of Postsecondary Education (FIPSE) to test the LibQUAL+™ instrument in a broad range of institutions.

Begun with only 12 ARL libraries in 1999, the LibQUAL+™ project has now reached the end of its FIPSE grant and is international in scope with just over 300 libraries participating in the spring 2003 survey. Since the project began, over 250,000 individuals have responded to the survey, creating the richest dataset of information on user perception of library service quality that exists. Although most of the 2003 participants were academic libraries, specialized libraries such as health sciences and law also participated. Consortia such as OhioLINK, the NY3Rs, and the Network of Alabama Academic Libraries (NAAL) have provided financial support for their members and received reports that analyze the survey data for consortia. For the first time in 2003, the protocol was applied to a set of public libraries in New York; a British English version was tested with United Kingdom postsecondary libraries, and a Canadian French translation of the instrument was prepared to support the participation of French Canadians. Results from this year's project are summarized in an article on page 8. After 2003, LibQUAL+™ will evolve into an ongoing service quality assessment program at ARL.

Additional funding from the National Science Foundation has been secured to adapt the LibQUAL+™ instrument for assessing digital library service quality within the National Science Digital Library (NSDL) framework. Qualitative re-grounding is ongoing with site visits, interviews, and focus groups held at the Math Forum, DLESE (Digital Library for Earth System Education), and MERLOT (Multimedia Educational Research for Learning and Online Teaching) to date.[8] Developing an understanding of the dimensions of digital library service quality is surfacing questions related to electronic services, quality, and educational impact, an area where relatively little theoretical and practical understanding exists to date. A report on the evaluation work of the NSDL appears on page 17.

Ease and Breadth of Access

The increasing use of electronic resources, documented by expenditures reported in ARL Supplementary Statistics and confirmed by LibQUAL+™ findings each year, was another area identified by retreat participants as a high priority for new measures. In February 2004, another set of ARL members began the E-Metrics project, an effort to explore the feasibility of defining and collecting data on the use and value of electronic resources. ARL had some experience in tracking expenditures for electronic resources through the ARL Supplementary Statistics but there was recognition that much more work was needed. The E-Metrics project was coordinated by Project Co-Chairs Sherrie Schmidt (Arizona State) and Rush Miller (Pittsburgh). ARL contracted with Florida State University's Information Use Management and Policy Institute to direct a study, gather data, and do the necessary analysis. The project was conducted in three phases: (1) an inventory of what libraries were already doing in this area and identification of any libraries that could provide best practices; (2) identification and testing of data elements that could be collected and used as measures of electronic resources for trend analysis and benchmarking; and (3) analysis of the connection between the use of electronic resources and institutional outcomes. Using surveys, consultations, field tests, and site visits, the project investigators identified 16 measures in four categories (resources, use, expenditures, and digitization) along with three performance indicators that ARL libraries might use to track electronic resource use. E-Metrics project documents can be found on the project's Web site[9] and were also published in print.

A related ARL Task Force on Statistics from Vendor-Based Database Products focused on working with vendors to arrive at agreed-upon data element definitions and terms, specific data that can be collected, and methods for reporting data to libraries. This task force built on existing work in the field. ARL also participated in international efforts culminating in early sponsorship of Project COUNTER (Counting Online Usage of NeTworked Electronic Resources), an international initiative designed to serve librarians, publishers, and intermediaries by facilitating the

recording and exchange of online usage statistics.[10] Release 1 of the COUNTER Code of Practice, which focuses on journals and databases, was agreed to and issued in December 2002.

After completion of the E-Metrics study, ARL participating libraries decided to continue further investigation and tested the proposed data elements again in 2002. Gordon Fretwell (University of Massachusetts) served as a Visiting Program Officer during the first half of the year and Martha Kyrillidou completed the data collection in the latter half. This testing experience led to the decision to open the 2003–04 pilot process to additional participants. The data elements collected as part of this project then will move into the regular ARL Supplementary Statistics collection cycle for the year starting in July 2004. Most of these items relate to the number and use of electronic resources and also seek to describe emerging digital library operations.

Library Impact on Teaching and Learning

For many institutions, the accreditation process has provided the strongest impetus for a shifting campus emphasis on institutional outcomes. The retreat participants recognized that some of the more difficult measures to develop would identify how a library contributes to the teaching and learning process. In order to address this issue, several ARL members provided funds to contract with Eller Distinguished Service Professor Kenneth Smith (University of Arizona) to prepare a white paper, "New Roles and Responsibilities for the University Library: Advancing Student Learning through Outcomes Assessment."[11] Smith's paper encourages libraries to develop strategies for becoming involved in campus assessment activities. He encouraged libraries to:

- Move from a content view (books, subject knowledge) to a competency view (what students are able to do)
- Understand learning outcomes of academic degree programs
- Develop curriculum segments or "offerings" through which the library achieves learning outcomes.[12]

ARL subsequently established a Learning Outcomes Working Group under the leadership of Carla Stoffle to expand on Smith's recommendations and identify measures that would describe the library's contribution to student learning outcomes. Several tasks were put forward to the group: survey accreditation agencies to identify best practice institutions, compile learning outcomes used by ARL institutions for specific academic departments or general education requirements (if they exist), contact professional associations to identify those that are working in this area, identify specific offerings that libraries could develop to meet outcomes thus far identified, establish a plan to take the offerings to faculty in a select group of institutions, explore with ACRL the offering of training on information literacy skills, define and create generalized tools for assessment, and, perhaps most importantly, define what is meant by learning outcomes and at what level ARL should focus its attention (student, course, program, department, etc.).

The working group began by scanning the higher education environment for consensus on or best practices in identifying learning outcomes, without much success. The results confirmed the importance, but also the difficulties, of advancing their agenda:

- National organizations had not yet agreed on a standard definition of learning outcomes.
- Accrediting agencies were leaving the discussion of learning outcomes up to the individual campuses.
- Since definition of learning outcomes is vague, most campuses have not yet grappled with the issue—those that have are doing it because of the pressure of accreditation.
- Identification of the people within an organization concerned with learning outcomes is difficult.
- Compilation of a common list of learning outcomes from ARL members is not yet possible.
- Best practices have been limited to smaller institutions—none are of the size and complexity of large research organizations.
- Faculty feel threatened by campus discussions of learning outcomes.

Fortunately, within the past couple of years, some of these challenges are being addressed. Recent literature on higher education assessment illustrates evidence that general education outcomes and, in some cases, outcomes for some professional programs are being developed and used. Common themes and learning objectives have arisen that can be applied by institutions beginning to develop their own outcomes assessment efforts. Accrediting agencies are a key source in encouraging these developments. The working group is preparing a white paper that summarizes the literature on higher education assessment, examines the definitions and sources of learning outcomes, identifies domains of learning, describes the levels of assessment, and suggests frameworks for putting these pieces together within an institution. See the related article on page 14 of this issue for a summary of the white paper.

One specific recommendation to discover libraries' contributions to the teaching and learning process is to begin with a process over which libraries have control, such as instruction, information literacy, or reference. ARL has taken that approach and is now partnering with Kent State University on Project SAILS (Standardized Assessment of Information Literacy Skills), which has received partial funding from the Institute of Museum and Library Services (IMLS) to develop an instrument for programmatic-level assessment of information literacy skills. SAILS is based on outcomes defined by the ACRL *Information Literacy Competency Standards for Higher Education*.[13] Seven libraries participated in Phase I of SAILS and met at the ALA Annual Conference in June 2003 to review their test results, suggest how data could be best formatted for use by libraries, and share their experiences in administering the survey to different student groups. Over forty libraries are participating in Phase II of the project, which will continue to test both the items in the survey as well as how best to administer the instrument within different library and academic settings. A fuller description of the SAILS Project appears on page 18.

Library Impact on Research

In 2001, Douglas Jones (University of Arizona) was appointed ARL Visiting Program Officer to identify issues and explore possible measures that demonstrate the role of the library in support of the university research mission. He investigated a set of dimensions that might lead to library success in support of the research mission of the university. His focus was on correlating research outcomes with different types of library support, identification of best practices at those institutions for which comparable levels of library support resulted in greater productivity, and interviews with key faculty and administrators to gather insights into how libraries might improve support for research. Some of the sources for institutional rankings and performance measures that might be used as indicators for research outcomes included *U.S. News and World Report* rankings, The Center (University of Florida) Top American Research Universities, the Association of American Universities Data Exchange (AAUDE) Data Warehousing Initiative, and the Integrated Postsecondary Education Data System (IPEDS).[14] In a presentation to the ARL survey coordinators in January 2002, Jones identified some possible research-related outcome measures: (1) scale of research (e.g., total R&D expenditures and federally sponsored R&D); (2) faculty (e.g., membership in national academies or other significant faculty awards indicating faculty distinction); (3) students (e.g., Ph.D.s awarded, post-doctoral appointments); (4) scholarly output (e.g., articles, books). Jones noted that it is difficult to isolate the library's contribution to research quality or productivity. There may be correlation between research outcome and library support, but causation is more difficult to prove. He suggested that libraries identify their institution's most productive researchers and ensure that their information needs are met.

If researcher access to information in a networked environment contributes to their success, finding a way to capture data on use of resources is one way that a library can document its contributions. Brinley Franklin (University of Connecticut) and Terry Plum (Simmons Graduate School of Library and Information Sciences) conducted patron use surveys of networked electronic services at five geographically disparate academic health science libraries and two main libraries between 1999 and 2003. They developed a pop-up Web-survey methodology that can be exploited to identify not only how library patrons use electronic services, but also to what purpose that use was being made. This study, called Measuring the Impact of Networked Electronic Services (MINES), is becoming an ARL new measures project. See the article on page 20.

Cost Effectiveness of Library Operations and Services

Libraries continually seek ways to improve the cost-effectiveness of their operations. Retreat participants identified this area as one that needed to be addressed, but did not identify any specific operations to be explored. There was general agreement that the 1997 ILL/DD Performance Measures study[15] resulted in significant changes in the operations of interlibrary loan departments across the membership. Members of the Statistics and Measurement Committee were interested in a follow-up study to see whether departments had become more efficient and whether costs had been reduced or contained. A new study, Assessing ILL/DD Services, was begun in fall 2002. This self-funded study updated, replicated, and expanded the 1997 ARL ILL/DD Performance Measures Study and obtained data on the performance of mediated and user-initiated (unmediated) interlibrary loan/document delivery operations in research, academic, and special libraries. A total of 75 ARL and non-ARL libraries participated. Tom Delaney (Colorado State University) and Bruce Kingma (Syracuse University) collaborated with Mary Jackson (ARL) in the study. See page 21 for an article on this project. ARL will publish a final report and sponsor a series of workshops to disseminate the findings.

The Statistics and Measurement program also investigated the Technical Services Cost Study conducted by five universities across the country. Using a time- and cost-recording process developed by Dilys Morris (Chronos Group, formerly of Iowa State University), staff time on tasks performed within

different cost centers is tracked over several sampled time periods. A variety of work processes can be traced and the resulting data can be analyzed by employee type, time, non-staff costs (such as vendors, systems, utilities), etc. Software developed by this project will enable libraries to input their own data and use it for benchmarking against the collective results.

Along the same lines, Eileen Hitchingham (Virginia Tech), with other members of the Statistics and Measurement Committee, worked on a Staff Allocation Project to determine if a set of spreadsheets could be developed into a tool for library managers to determine how staff time was being distributed across a range of functions. Two institutions, Notre Dame and Vanderbilt, will be combining the methodologies from the Technical Services Cost Study and the Staff Allocation Project to look at staff allocation costs across the library. See the article by Dilys Morris on page 23.

In order to change the culture, change what you measure.

Library Facilities and Space

While no specific projects have begun yet regarding measurements of library space, this area has not languished. ARL and OCLC cosponsored a Strategic Issues Forum, "Future Library Architecture: Conception, Design, and Use of Library Space" in February 2002 in Las Vegas, Nevada. The goals of the forum were (1) to engage library leaders in an exploration of important issues affecting the future development of library space and (2) to help prepare them to envision and make the case for innovative library spaces that are directly responsive to user needs and interests on campus. In addition, ARL's May 2003 Membership Meeting examined the connections between the concepts of research libraries as place and as space and how they contribute to a sense of community.[16]

One question retreat participants asked about space was, "How can libraries tell if they are effectively utilizing their space?" Cliff Haka and Jim Hensley (Michigan State University), in an article on page 24, describe Michigan State University's effort to conduct a space-utilization study using CAD 2000 software. They postulate that this might be a useful tool for other libraries and could perhaps provide an opportunity for benchmarking the use of space in ARL libraries.

Organizational Capacity

Retreat participants acknowledged that one of the challenges facing research libraries is the importance of developing an institutional culture that understands the need to focus on new measures (i.e., outcomes) rather than solely on traditional measures. One way to change culture is to change what is measured. This shift in

thinking requires library staff to not only learn new tools, but also to recognize the importance of data in making decisions that will improve library services and operations. ARL and the Office of Leadership and Management Services have developed a number of activities to support the learning needed by library staff. The LibQUAL+™ program has a range of training, including introductory workshops, project-planning sessions, data-review sessions, and post-survey seminars. The program also offers the Service Quality Evaluation Academy, an intensive five-day workshop focusing on qualitative and quantitative methods for collecting and analyzing library service quality data. In collaboration with the ARL/OLMS Online Lyceum, a Measuring Library Service Quality Course has been offered to all LibQUAL+™ participants and a range of libraries interested in the background and theory of measuring service quality and methods to assess and improve service. The ARL/OLMS offers a two-day workshop on Creating a Culture of Assessment, which identifies the main forces influencing current library goals and work environments and how these forces demand the adoption of a new paradigm of thinking, new attitudes, and the support for structures that make assessment and evaluation an integral part of library work. See page 31 for more information.

"Toward Rethinking Academic Library Performance in the Digital Age," an ARL/OCLC Strategic Issues Forum held in 2001, introduced senior library administrators to issues regarding performance measurement, and provided an overview of the Balanced Scorecard methodology as one means for libraries to establish performance management systems. The University of Virginia is one of the first libraries to use this methodology; an article by Jim Self describing their experience is on page 28.

In 1998, the University of Arizona Library implemented a performance measurement system to assess itself. An article on page 26 describes Arizona's experience with the Performance Effectiveness Management System that aligns individual, unit, and whole organization efforts with library strategic goals and provides measures that indicate success, progress, and the need for improvement. The article also gives examples of benchmarks that Arizona identified in order to provide other libraries with a point of departure for measuring their own operations.

Susan Beck (Rutgers University) served as an ARL Visiting Program Officer to investigate the impact of assessment on library management decision making and the degree to which assessment data has influenced organizational change. She studied nine public ARL institutions, selected primarily because they are involved in assessment activities. See the article on page 30 for a summary of her research.

Moving Measures into the Mainstream

What does it take for an institution to move into a culture of assessment? Based on experience and research over the last few years, the following have emerged as essential elements:

- Resources (i.e., 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

The ARL New Measures Initiative has worked to assure that all of these components are part of the projects undertaken. After several years of hard work testing and refining measures and developing a broader base of experience and expertise among staff in libraries, it is now time to begin moving many of the "new" measures into the mainstream of library operations.

Moving new measures into the mainstream of the ARL Statistics and Measurement program does not mean ARL is throwing the baby out with the bathwater. The data ARL traditionally gathers are still important for documenting the transformation of research library collections, services, and staff, and for identifying trends in the costs of operations. However, the changing nature of research libraries also calls for new and different measures that not only describe the library, but also indicate how well a library is serving its community.

There is no doubt that demands for library accountability will continue. The users of library services have an insatiable appetite for quality information anytime, anywhere. Parent institutions are pressured to demonstrate their value to legislators, funding authorities, students, etc., and are asking for data from all organizational units, including libraries, in order to provide it. An exploding growth in the use and applications of technology along with an increased competition of resources makes it more important than ever for libraries to use hard data to make the case for resources. This means libraries will have to continually measure and monitor their effectiveness and, as technology and user behavior changes, strive to develop measures that have the greatest potential in helping library managers lead change.

The Success of the New Measures Initiative

There are several characteristics of success of the ARL New Measures Initiative. The first has been the commitment by individual institutions to take leadership for specific projects. As highlighted in this article, individual institutions and library directors have taken leadership for new measures projects. Their personal interest and dedication to the initiative itself or to a particular area for investigation cannot be overstated. They have ensured not only that projects are well defined and launched, but also have kept the enthusiasm for the work in the forefront of the ARL community. Another characteristic has been the project nature of the initiative. As agreed to by retreat participants, it was evident that no single measure or set of measures would work to address all eight areas of concern. By allowing a variety of projects and approaches, the initiative was able to quickly address all of the areas relative to the degree of interest expressed by individual project participants. Also, member libraries were not obligated to join all projects. They could choose only those that were of most interest based on local situations and those choices guaranteed their interest and active participation. Joining a project could mean project leadership, active participation by the director and staff, or support through funding. As results of each of the project are made available, the full ARL community as well as the larger library community benefits. Another aspect of success for these projects is the opening up of some of them to the wider library community. The cross-fertilization especially with other academic institutions has allowed the projects to expand their support base and strengthened the acceptance of the tools and methodologies throughout the library community.

The data that libraries collected and reported in the past were primarily input data (how much do you have). In the 1990s there was encouragement to provide more output data (how much do you do). Libraries are now in an environment that asks for outcomes (what is the benefit of your service) or impact (what difference do you make). Institutions are identifying their outcomes, defining criteria for success, and determining methods to look for evidence of success. ARL has demonstrated that it is possible to move from input to impact measures through its New Measures Initiative. As the initiative becomes a regular component of the Statistics and Measurement program, its success factors can serve as elements in a model for other organizational units to use as they look to improve their own institutional effectiveness.

The year 2003 is transitional for the New Measures Initiative—leadership of the Statistics and Measurement Committee passed from Carla Stoffle to Brinley Franklin. The initiative has matured and been integrated into the Statistics and Measurement program. For example, ARL has assumed full responsibility for LibQUAL+™, plans are in place for E-Metrics to move into the ARL

Supplementary Statistics, Project COUNTER is being implemented, and results of the new ILL/DD study will be used by libraries to improve their operations. We have reached a point where some measures that have been successfully tested will be moved to the main data-collection program while others are dropped from ARL's surveys. Other measures will continue to be tested among a larger community in a continuous cycle of evaluate and deploy or drop. Ideas for new measures continue to be welcomed by the program. Several years ago our challenge was to develop our own new measures or they would be determined by others for us. We have met that challenge successfully through the New Measures Initiative.

1 Sarah Pritchard, "New Directions for ARL Statistics," *ARL Bimonthly Report*, no. 161; Martha Kyrillidou and William Crowe, "In Search of New Measures," *ARL Bimonthly Report*, no. 197 (April 1998), <http://www.arl.org/newsltr/197/newmeas.html>.

2 ARL characterized some of the issues surrounding measurement in research libraries in a special issue of its newsletter, "Issues in Research Library Measurement," *ARL Bimonthly Report*, no. 197 (April 1998), <http://www.arl.org/newsltr/197/197toc.html>.

3 For those readers not familiar with the development of the New Measures Initiative, it was introduced in the December 1999 *ARL Bimonthly Report*, no. 207, <http://www.arl.org/newsltr/207/>.

4 <http://www.arl.org/stats/newmeas/nmbackground.html>.

5 <http://www.libqual.org/>.

6 <http://www.libqual.org/Publications/>.

7 "Measuring Service Quality," Martha Kyrillidou and Fred M. Heath, eds. *Library Trends* 49, no. 4: 541-799, <http://www.arl.org/libqual/events/oct2000msq/program.html>.

8 <http://mathforum.org/>, <http://www.dlese.org/dds/>, <http://www.merlot.org/Home.po>.

9 <http://www.arl.org/stats/newmeas/emetrics/>.

10 <http://www.projectcounter.org/>.

11 <http://www.arl.org/stats/newmeas/outcomes/HEOSmith.html>.

12 Examples of how research libraries are engaged in developing curriculum segments and offering instructional technology services are available in the responses to the 2002 ARL Survey on New Approaches to Collections and Access services, <http://db.arl.org/Casort/>.

13 *Information Literacy Competency Standards for Higher Education,* approved by the Association of College and Research Libraries in January 2000, <http://www.ala.org/Content/NavigationMenu/ACRL/Standards_and_Guidelines/Information_Literacy_Competency_Standards_for_Higher_Education.htm>.

14 <http://www.usnews.com/usnews/edu/college/rankings/rankindex_brief.php>, <http://thecenter.ufl.edu/>, <http://www.pb.uillinois.edu/AAUDE/>, <http://nces.ed.gov/ipeds/>.

15 Mary Jackson, *Measuring the Performance of Interlibrary Loan Operations in North American Research & College Libraries* (Washington, D.C.: ARL, 1998).

16 <http://www.arl.org/arl/proceedings/142/>.

MEASURING USER SATISFACTION AND MARKET PENETRATION

LIBRARY USERS ASSESS SERVICE QUALITY WITH LIBQUAL+™ AND E-QUAL

by Kaylyn Hipps, Editorial and Research Associate, ARL, and Martha Kyrillidou, Senior Program Officer for Statistics and Measurement, ARL

LibQUAL+™ is a suite of services aimed at measuring library user perceptions and expectations of service quality. At the core of the project is a Web-based survey that seeks feedback from library users about their experiences with the library and the extent to which the library meets their needs. The survey was developed by a team from ARL and Texas A&M University through a combination of qualitative interviews with North American library users and four annual survey iterations to date, starting in 2000. The U.S. Department of Education's Fund for the Improvement of Postsecondary Education (FIPSE) provided partial support for the first three years of the project.

After four iterations and revisions of the survey instrument, three dimensions of library service quality have surfaced:

- library as place (utilitarian space, symbol, refuge);
- information control (scope, timeliness, convenience, ease of navigation, modern equipment); and
- affect of service (empathy, responsiveness, assurance, reliability).

In the spring of 2003, more than 120,000 library users from over 300 libraries responded to the LibQUAL+™ survey. These numbers include over 25,000 users at 55 ARL academic libraries, who said, in the aggregate, that their libraries were most successful in providing physical facilities that meet their needs and in providing skilled and caring staff. These users expressed the highest expectations for those services and tools that enable them to easily gain access to and use information independently; these expectations are part of the library service quality dimension called "information control." The area that ARL academic library users identified as especially needing improvement is access to information that they need for their work, such as journal collections and other materials in both print and electronic form. This is also part of the information control dimension.

Participating institutions are using their LibQUAL+™ survey data to implement change within their libraries. Many of the libraries already knew that some changes were necessary; the survey confirms those beliefs and provides evidence to justify targeted funding increases. Some libraries are using additional data-gathering methods—such as focus groups and local surveys—to supplement their LibQUAL+™ results or investigate them in more detail. It is also common for libraries who have participated in LibQUAL+™ to strengthen their communications with users in order to better inform them and manage their expectations. Libraries find the user comments that are gathered as part of the survey extremely helpful as they interpret the results from the standardized part of the survey in the context of their local environments.[1]

The LibQUAL+™ data set is by far the largest of its kind. Each annual survey iteration has seen tremendous growth in the numbers of participating libraries and survey respondents as well as a broadening of the types of libraries. The number of participating institutions increased from 12 in 2000 to 308 in 2003. The 2003 participants included college and university libraries, community college libraries, health sciences libraries, military libraries, public libraries, and state libraries. The survey was completed by more than 5,000 users in 2000 and more than 120,000 in 2003. Libraries in four countries—the U.S., Canada, the U.K., and the Netherlands—participated in the 2003 survey.

For the first time, in 2003, the LibQUAL+™ survey was translated into languages other than American English: a British English version was administered in 20 U.K. institutions, a Dutch English version in one institution in the Netherlands, and a Canadian French version in two universities in Canada. An analysis of the survey data from the French Canadian institutions showed that the translation process produced scores that are equivalent to the English versions of the instrument. The LibQUAL+™ team is confident that the different versions of the survey are "culturally relevant" in their respective languages and "conceptually equivalent to the original" American English survey.[2]

LibQUAL+™ is now offered as an ongoing ARL program to individual libraries of all types. In addition, many libraries have participated in the program as part of consortia. Registration for the spring 2004 survey is underway this fall.

e-QUAL
The most recent development in the LibQUAL+™ project is the team's adaptation of the survey to the digital library environment. This effort, called e-QUAL, is partially supported by a grant from the National Science Foundation, National Science Digital Library (NSF/NSDL). The overarching goal of e-QUAL is to develop a digital library service quality assessment process that enhances student learning by permitting the allocation of resources to areas of user-identified need. The e-QUAL project seeks to accomplish this goal by achieving the following objectives and outcomes:

- define the dimensions of digital library service quality from the perspective of the users;
- develop a tool for measuring user perceptions and expectations of digital library service quality across NSDL digital library contexts;
- identify digital library "best practices" that permit generalizations across operations and development platforms;
- enhance student learning by effectively managing user perceptions and expectations of digital library services;
- establish a digital library service quality assessment program as an integral part of the library service quality assessment program at ARL; and
- institutionalize continuous product and process evaluation efforts directed toward positive and timely management of outcomes.[3]

Many digital libraries are still in their formative stages of development so the creation of an evaluation tool like the envisioned e-QUAL survey needs to be undertaken with care. To avoid premature evaluation and misguided comparisons across very different digital libraries, the e-QUAL project advocates using mixed methods to develop a digital library assessment tool and anticipates that e-QUAL eventually will be one of several such tools for digital libraries.

The e-QUAL project is currently in the qualitative stage of research. The team is studying digital libraries—such as the Math Forum, MERLOT, and DLESE—via interviews, site visits, and focus groups. The qualitative analysis will help the project understand digital library users' perceptions of service quality.

Although there are different working models of digital libraries—e.g., some contain actual collected materials, posted online; others contain links to materials hosted elsewhere; others are developing infrastructure tools—they do share the goal of connecting users with information via the Internet. Characteristics of digital libraries that may be evaluated include information access, sense of community, and quality of materials.

The e-QUAL project will develop and test an initial set of survey questions in late 2003. A second round of testing is planned for spring 2004.

~~~

The LibQUAL+™ and e-QUAL team members include:

Consuella Askew, LibQUAL+™ Program Specialist, ARL

Julia Blixrud, Assistant Executive Director, External Relations, ARL

Colleen Cook, Interim Dean of University Libraries and Professor and Wright Chair of Library Science, Texas A&M University

Fred Heath, Vice Provost and Director of General Libraries, University of Texas at Austin

Amy Hoseth, New Measures Project Assistant, ARL

Martha Kyrillidou, Senior Program Officer for Statistics and Measurement, ARL

Yvonna Lincoln, Distinguished Professor and Ruth Harrington Chair of Education, Educational Administration, Texas A&M University

Jonathan Sousa, Technical Applications Development Manager for New Measures, ARL

Bruce Thompson, Professor and Distinguished Research Scholar, Educational Psychology, Texas A&M University

Duane Webster, Executive Director, ARL

Mark Young, Research Assistant for Statistics and Measurement, ARL

For more information about LibQUAL+™, see <http://www.libqual.org/>. For more about e-QUAL, <http://www.arl.org/stats/newmeas/emetrics/nsdl.html>.

[1] For more on how libraries use LibQUAL+™ data, see Consuella Askew Waller and Kaylyn Hipps, "Using LibQUAL+™ and Developing a Culture of Assessment in Libraries," *ARL Bimonthly Report*, no. 221 (April 2002): 10–11, <http://www.arl.org/newsltr/221/libqarl.html>; and Robyn Huff-Eibl and Shelley Phipps, "Using LibQUAL+™ Results at the University of Arizona: Responding to Customer Input—Listening and Acting," *ARL Bimonthly Report*, no. 221 (April 2002): 12–13, <http://www.arl.org/newsltr/221/libqaz.html>.

[2] Martha Kyrillidou, Toni Olshen, Fred Heath, Claude Bonnelly, and Jean-Pierre Cote, "Cross-Cultural Implementation of LibQUAL+™: The French Language Experience," presented at the Fifth Northumbria International Conference on Performance Measurement in Libraries, Durham, England, July 28, 2003.

[3] Colleen Cook, Fred Heath, Martha Kyrillidou, Yvonna Lincoln, Bruce Thompson, and Duane Webster, "Developing a National Science Digital Library (NSDL) LibQUAL+™ Protocol: An E-Service for Assessing the Library of the 21st Century," presented at the NSDL Evaluation workshop, Washington, D.C., October 2003, <http://www.libqual.org/documents/admin/NSDL_workshop_web.pdf>.

A LibQUAL+™ Primer

Excerpted from Colleen Cook, Fred Heath, and Bruce Thompson, "LibQUAL+™ from the U.K. Perspective," presented at the Fifth Northumbria International Conference on Performance Measurement in Libraries, Durham, England, July 28, 2003, <http://www.libqual.org/documents/admin/northumb_CC.pdf>.

LibQUAL+™ is only one of 11 different ways to listen to customers, called a "total market survey."[1] Berry recommended using multiple listening methods, and emphasized that "Ongoing data collection…is a necessity. Transactional survey, total market surveys, and employee research should always be included."

LibQUAL+™ was modeled on the 22-item SERVQUAL tool developed by Parasuraman, Berry and Zeithaml.[2] …However, SERVQUAL has been shown to measure some issues not particularly relevant in libraries…, and to not measure some issues of considerable interest to library users.

The final 22 LibQUAL+™ items were developed through several iterations of quantitative studies involving a larger pool of 56 items. These 56 items were themselves identified following qualitative research interviews with library student and faculty users at several different universities.

In conducting service quality assessments, it is absolutely critical to demonstrate that the assessment scores measure something (as against nothing). Scores measure nothing when they are random. In psychometrics, the property of data measuring something is called "reliability." In previous administrations, LibQUAL+™ scores have been repeatedly shown to be reliable.

Even when assessment scores are demonstrated to be reliable, it additionally must be shown that the scores measure the intended constructs. [This] is called "validity." A primary tool in such proofs involves the statistical method called factor analysis. Various methods have been employed in prior studies suggesting that LibQUAL+™ scores are valid.

There are two primary ways to interpret LibQUAL+™. First, scores on perceptions may be compared against scores on what is reported to be minimally acceptable service, and what is reported to be desired service; this is called the "zones of tolerance" interpretation framework.

Second, statistical norms may be used to characterize factually what percentage of users or of institutions generated lower perception ratings.

—Copyright 2003 Colleen Cook, Fred Heath, and Bruce Thompson

[1] L.L. Berry, *On Great Service: A Framework for Action* (New York: The Free Press, 1995).

[2] A. Parasuraman, L.L. Berry, and V.A. Zeithaml, "Refinement and Reassessment of the SERVQUAL Scale," *Journal of Retailing* 67 (1991): 420–450.

E-METRICS: NEXT STEPS FOR MEASURING ELECTRONIC RESOURCES

by Julia C. Blixrud, Assistant Executive Director, External Relations, ARL, and Martha Kyrillidou, Senior Program Officer for Statistics and Measurement, ARL

Responding to user demand, libraries have steadily been shifting the focus of their collection development to the acquisition and licensing of electronic content, much of it via consortia. In the last decade, the average percentage of acquisition dollars that ARL member libraries directed to electronic resources rose from 3% to 20%. In 2002, 110 ARL university libraries reported spending more than $171 million on electronic resources, and 48 ARL libraries reported another $20 million expended on their behalf through centrally funded consortia.[1] In spite of considerable efforts, aggregate data documenting the quantity and use of these e-resources was proving to be elusive. The ARL E-Metrics work was established in 2000 to develop standard definitions for measures that libraries could use to describe: (a) the e-resources they make accessible, (b) the use made of the resources, and (c) the level of library expenditures.

Background

ARL began collecting data on electronic resources through its Supplementary Statistics survey in 1992–93.[2] Many libraries were able to report some initial cost data but after several attempts the libraries were unable to count consistently the resources purchased or the use made of those resources. In 1996–98, Tim Jewell (University of Washington) served as an ARL Visiting Program Officer to analyze the Supplementary Statistics survey data on electronic resources from 1992–93 to 1995–96. His project confirmed that reported figures underestimated the actual expenditures by an unknown amount and that local record-keeping practices made it difficult to respond to some questions in the survey. The 1996–97 ARL Supplementary Statistics questionnaire was revised to try to capture figures that had been elusive in previous surveys. Jewell identified several other trends in the data that member libraries were reporting and confirmed that existing data was neither comprehensive nor comparable.[3]

Recognizing the pace at which e-resource spending was increasing, it was clear that consistent and reliable data was needed to identify whether the significant investment in these resources was benefiting library users. In 1999, the participants at the ARL New Measures retreat also identified electronic resources as an area of

continuing concern and in February 2000 a group of ARL directors and senior staff met for two days to determine how to tackle the complicated issue of gathering data on e-resources. A number of issues were identified at that meeting, many reiterating the points in the Jewell report. Some individuals noted they still had little data. For those libraries that had data, the data were inconsistent and unreliable. The libraries themselves had a variety of internal structures and procedures surrounding the acquisition and deployment of electronic resources and were not well organized to collect common data. Another major problem identified was the dearth of information about how the use of electronic information resources contributed to library user success.

Meeting participants agreed to contribute funds to address this issue and the ARL E-Metrics project was launched with 24 libraries participating. Led by Rush Miller (University of Pittsburgh) and Sherrie Schmidt (Arizona State University), the initial project began in May 2000 and was completed in December 2001. Florida State University's (FSU) Information Use Management and Policy Institute was engaged to conduct the study. It was completed in three phases: (1) an inventory of what libraries were already doing in this area and identification of any libraries that could provide best practices; (2) identification and testing of data elements that could be collected and used as measures of electronic resources for trend analysis and benchmarking; and (3) analysis of the connection between the use of electronic resources and institutional outcomes. The work of this period has been described extensively in the published literature.[4] A key outcome of the FSU Institute work was a set of recommended statistics and measures (see accompanying box). In keeping with the goals of the ARL New Measures Initiative, documents from the E-Metrics project are available to the wider library community.[5]

During the course of this project, a Working Group on Database Vendor Statistics was established to investigate how to effectively gather and report data from vendor databases since much of the e-resource measure counts would come from those sources. ARL built relationships with other parties interested in vendor statistics as efforts moved forward both nationally and internationally. ARL continues to work with the vendor community, primarily by supporting Project COUNTER (Counting Online Usage of NeTworked Electronic Resources), a multi-agency project developing a single, internationally accepted Code of Practice that will allow the usage of online information products to be measured more effectively.[6] Project COUNTER is now soliciting membership support from libraries as well as other constituencies. ARL also monitors developments within the national and international standards communities as NISO and ISO draft standards for measuring electronic resources are finalized.[7]

11

E-METRICS CHRONOLOGY

1993 82 libraries report e-resources data for ARL Supplementary Statistics

1996–98 Assessment of ARL Supplementary Statistics survey data on e-resources undertaken by Tim Jewell (University of Washington) as ARL Visiting Program Officer

1999 ARL New Measures Retreat, e-resources identified as one of the key areas of concern

105 libraries report e-resources for ARL Supplementary Statistics

2000 Two-day meeting of library directors and senior staff focuses on gathering data on e-resources

ARL E-Metrics Project launched with 24 libraries participating; conducted in three phases over two years by Florida State University's (FSU) Information Use Management and Policy Institute

Working Group on Vendor Statistics established

2001 FSU Project team recommends 16 measures of e-resources and 3 measures of the use or performance of these resources.

ARL joins Project COUNTER

2002 Call for participation to ARL libraries to test the measures recommended by FSU team; 35 libraries respond and commit to collect the data for 2001–02 or part of the 2002–03 academic year and compare results with vendor data

Gordon Fretwell (University of Massachusetts), Visiting Program Officer, and Martha Kyrillidou (ARL) collect and analyze data

Project COUNTER issues draft guidelines for measuring usage of online information products

2003 Decision reached that data being collected related to expenditures for e-resources is stable enough to move from Supplementary Statistics to the main ARL Statistics as of the 2003–04 collection cycle (July 2004)

Decision reached to move all other E-Metrics data elements still being tested into the Supplementary Statistics beginning with the 2003–04 collection cycle (July 2004)

Call for participation in the 2003–04 data collection of E-metrics issued

E-Metrics Pilot Projects

Acceptance of the E-Metrics project report from the FSU Institute was one step in the series of ARL activities addressing the issue of measuring electronic resources. Next came a call for participation in the summer of 2002 for ARL member libraries to continue the work of the project and test the recommended measures in an E-Metrics Pilot Project during 2002–03. Thirty-five institutions responded, half of them new to the project

For the 2002–03 E-Metrics Pilot Project, the participating libraries were asked to report back through a structured questionnaire to Gordon Fretwell (University of Massachusetts), who served as an ARL Visiting Program Officer. The questionnaire gathered information on what data the library was currently collecting, how much of it could be reported for FY 2001–02, any additional statistics being collected on in-house electronic resources, and what types of internal analysis were being done with the data collected. Participants were also asked to report data for the first quarter of the fiscal year 2002–03 (July 2002 through September 2003). The goal was to collect some in-depth annual data from some vendors at some libraries, and a wider collection of data for a briefer window of activity (first quarter of current fiscal year) from a larger number of libraries. Variations in vendor reporting may mean that data initially examined would not be strictly comparable across vendors.

An in-person project meeting at the January 2003 ALA Midwinter Meeting provided an opportunity for some libraries to present and discuss the analysis techniques and specific output or products that they have developed and utilized.[8] As the year progressed, a number of the 2002–03 E-Metrics participants spent time organizing the collection of "database use" statistics (sessions, searches, items). The redesigned ARL E-Metrics Web page includes a section where participating libraries' spreadsheets can be seen as examples of the new E-Metrics data-collection process.[9]

Current Status of Pilot Project and Next Steps

In light of their proven utility and reliability, several data elements for electronic resources previously collected through the ARL Supplementary Statistics will be moving into the regular ARL Statistics data-collection activity beginning in July 2004 with collection of data for 2003–04. They are items related to expenditures on electronic resources: Computer Files (one-time/monographic purchase); Electronic Serials; Bibliographic Utilities, Networks, and Consortia (Library); and Bibliographic Utilities, Networks, and Consortia (External). Items remaining in the ARL Supplementary Statistics that are continuing to be tested as part of the E-Metrics Project include Expenditures on Electronic indexes and reference tools and Expenditures on Electronic full-text periodicals.

E-Metrics Being Tested

Patron-Accessible Electronic Resources
- Number of electronic full-text journals
- Number of electronic reference sources
- Number of electronic books

Use of Networked Resources & Related Infrastructure
- Number of electronic reference transactions
- Number of logins (sessions) to electronic databases
- Number of queries (searches) in electronic databases
- Items requested in electronic databases
- Virtual visits to library's Web site and catalog

Expenditures for Networked Resources & Related Infrastructure
- Cost of electronic full-text journals
- Cost of electronic reference sources
- Cost of electronic books
- Library expenditures for bibliographic utilities, networks, & consortia
- External expenditures for bibliographic utilities, networks, & consortia

Library Digitization Activities
- Size of library digital collection
- Use of library digital collection
- Cost of digital collection construction & management

All the data elements collected through the current ARL E-Metrics pilot will be moving into the regular ARL Supplementary Statistics collection cycle starting in July 2004 for the year 2003–04. This move will begin to test more widely the data collection that thus far has been conducted on a pilot-project basis. During 2003–04, the pilot process was opened to additional participants.

Training

As with all of the new measures projects, training was identified by many E-Metrics participant libraries as a major need. In addition to providing specific information about how to gather the recommended statistics, training also should include conceptual understanding of the importance of using data for decision making and developing a culture of assessment that encourages staff to make use of gathered data. ARL provides opportunities for E-Metrics project participants to share experiences at meetings and delivers workshops about gathering and using data in a variety of ways.[10]

Conclusion

Increasingly libraries are called upon to demonstrate the educational impact and outcomes achieved because of the increased availability of electronic resources. In the new measures environment, we are faced with the challenge of identifying (a) new "species" of metrics and (b) new methods for measuring the effectiveness of these new species. The work described above primarily focuses on the identification of new species, i.e., new data categories that will give us some general sense of the evolution of electronic resources in terms of input and output measures. There are also efforts underway, such as the MINES Project described by Franklin and Plum in this issue and the effort to develop e-QUAL for the digital library environment, that are attempting to develop new user-based methods for understanding the value derived from the use of electronic resources. ARL continues to seek and develop the best measures to demonstrate how the provision of electronic resources contributes to library users' success.

1 *ARL Supplementary Statistics 2001–02* (Washington, D.C.: ARL, 2003).

2 A printed compilation of the investigator reports and the instruction manuals are available for purchase and all the documents are on the E-Metrics Web site <http://www.arl.org/stats/sup/>.

3 A full report was sent to each ARL member library as part of the report on the 1995–96 ARL Supplementary Statistics and is also available at <http://www.arl.org/stats/specproj/etrends.htm>.

4 Rush Miller and Sherrie Schmidt, "E-Metrics: Measures for Electronic Resources," in *Proceedings of the 4th Northumbria International Conference on Performance Measurement in Libraries and Information Services* (Washington, D.C.: ARL, 2002): 37–42; Wonsik "Jeff" Shim, Charles R. McClure, and John Carlo Bertot, "Preliminary Statistics and Measures for ARL Libraries to Describe Electronic Resources and Services," in *Proceedings of the 4th Northumbria International Conference on Performance Measurement in Libraries and Information Services* (Washington, D.C.: ARL, 2002): 337–344; Wonsik "Jeff" Shim, Charles R. McClure, and John Carlo Bertot, "Measures and Statistics for Research Library Networked Services: ARL E-Metrics Phase II Report," *ARL Bimonthly Report* 219 (Dec. 2001): 8–9, <http://www.arl.org/newsltr/219/emetrics.html>; Wonsik "Jeff" Shim, Charles R. McClure, and John Carlo Bertot, "Data Gathering Practices in the Networked Environment," *ARL Bimonthly Report* 213 (Dec. 2000): 6–8 <http://www.arl.org/newsltr/213/data.html>; Wonsik "Jeff" Shim and Charles R. McClure, "Measuring Services, Resources, Users, and Use in the Networked Environment," *ARL Bimonthly Report* 210 (June 2000): 9–10, <http://www.arl.org/newsltr/210/emetrics.html>; Julia C. Blixrud, "Measures for Electronic Use: The ARL E-Metrics Project," in *Statistics in Practice—Measuring and Managing: Proceedings of IFLA Satellite Conference*, Loughborough, England, August 2003 (Loughborough: Library & Information Statistics Unit (LISU), Loughborough University, 2003): 73–84, <http://www.lboro.ac.uk/departments/dis/lisu/statsinprac.html>.

5 <http://www.arl.org/stats/newmeas/emetrics/>.

6 <http://www.projectcounter.org/>.

7 NISO Z39.7-2002 Draft Standard for Trial Use: Information Services and Use: Metrics & Statistics for Libraries and Information Providers—Data Dictionary, <http://www.niso.org/emetrics/>.

8 The meeting was hosted by the University of Pennsylvania, which has been developing the Penn Library Data Farm, a repository of quantitative information to aid in the measurement and assessment of library resource use <http://metrics.library.upenn.edu/prototype/about/>.

9 Like every Web-redesign it is work in progress located at <http://www.arl.org/stats/newmeas/emetrics/>.

10 ARL is offering a Web cast in October 2003 on data-collection changes, <http://www.arl.org/arl/workshops.html>.

LIBRARY IMPACT ON TEACHING AND LEARNING

RESEARCH LIBRARY INVOLVEMENT IN LEARNING OUTCOMES ASSESSMENT PROGRAMS

by Sheila Young, Science Reference Librarian/Engineering, Arizona State University, and Julia C. Blixrud, Assistant Executive Director, External Relations, ARL

How could research libraries participate in the assessment of learning outcomes at the institutional level and determine how to measure their contribution to the learning process? This is the question now being pursued by the ARL Learning Outcomes Working Group. In 2000, Eller Distinguished Service Professor Kenneth Smith, University of Arizona, prepared a white paper for ARL that discussed how a library might begin developing shared learning outcomes with the larger campus community and then create "offerings" based on those shared outcomes, to integrate into the curriculum.[1] Dr. Smith recommended that libraries "understand the processes that are used to define learning outcomes, to select measures, to collaborate with other academic departments, and to use the results to improve their programs."

The Learning Outcomes Working Group, initially chaired by Carla Stoffle, Dean of Libraries at the University of Arizona, was established to follow up on Dr. Smith's recommendations. The working group is currently led by Karen Williams, Team Leader for Undergraduate Services, University of Arizona Library. As with other New Measures Initiative projects, the working group was charged to identify existing measures that could be built upon and applied in libraries. A scan of the higher education environment found that, while the emphasis on establishing learning outcomes and assessing those outcomes has gained momentum in recent years, campuses are still in various stages of grappling with this issue. Given this, the working group decided to develop their own white paper that would describe this environment and, using it as a foundation for understanding the current state of research and implementation in this field, decide how to take Dr. Smith's recommendations a step further.

The working group's white paper, now in draft, examines the definitions and sources of learning outcomes, domains of learning, levels of assessment, and frameworks suggested for putting these pieces together within an institution. What follows is a summary of the working group white paper.

Definitions of "Learning Outcomes"

Numerous authors express the concept of student learning outcomes as the changes that occur within the student, not what the instructor does.[2] That is, learning outcomes are "a level of knowledge, skills, abilities that a student has attained."[3] This emphasis on the change that takes place within the student is also reflected in a definition developed by Battersby and embraced by members of the working group.

> "Learning outcomes are the essential and enduring knowledge, abilities (skills), and attitudes (values, dispositions) that constitute the integrated learning needed by a graduate of a course or program."[4]

Sources of Learning Outcomes: National

Regional accrediting agencies are one source of the criteria upon which an institution will be evaluated. Accreditation criteria have been recently revised reflecting an increased emphasis on assessment and institutions are beginning to gain experience with these new expectations. Lindauer has examined and analyzed documents of the accreditation agencies and evaluates trends and specific text related to libraries. Her work presents examples from these documents and makes five observations and recommendations for library action. These concern the role of the library in information literacy, the role of the library as an "academic support unit," librarian involvement in accreditation standards, promotion on campus by librarians of the ACRL information literacy documents, and the need for librarians to stay informed about the current assessment projects of the ACRL, ARL, the Association of American Higher Education, and other organizations involved with assessment.[5]

Discipline-specific documents are an additional source of learning outcomes that are valued by departments across campuses. On a national level, the American Chemical Society, the American Psychological Association, and the Accreditation Board for Engineering and Technology Inc. have all developed such statements. The Council for Higher Education Accreditation (CHEA) lists a directory of "Specialized and Professional Accrediting Organizations" that are either recognized or will undergo a review by them.[6] The accreditation documents of these professional organizations periodically undergo revision and this may provide librarians an opportunity to contribute to the new standards.

Another source of learning outcomes from a national body is a 1995 report from the U.S. Department of Education, Office of Educational Research and Improvement: "National Assessment of College Student Learning: Identifying College Graduates' Essential Skills

in Writing, Speech and Listening, and Critical Thinking: Final Project Report."[7] The report identifies "higher order communication and thinking skills that college graduates should achieve" and describes the skills of analyzing, locating, and evaluating information as dimensions of critical thinking.

Within the library community, ACRL has produced information literacy and competency standards for higher education[8] and a standards tool kit[9] that lists the outcomes for each standard. In addition, an ACRL "Best Practices" initiative has developed "Characteristics of Programs of Information Literacy that Illustrate Best Practices: A Guideline," which was approved by the ACRL Board in 2003.[10] All of this work is useful as librarians examine the needs and challenges that are particular to their own institutions.

Sources of Learning Outcomes: Local
Statements by institutional governing boards indicate valued outcomes, including learning outcomes, if present. Examining the documents produced by the governing board, institution (e.g., the mission and the strategic plan), and the individual colleges, departments, and programs will indicate those learning outcomes that are of greatest importance.

Integration of Learning Outcomes through All Levels
Writing learning outcomes that are internally consistent with the department or discipline, the institution, state, and accrediting agency requires an examination of documents from all these levels. Huba describes this process of integration as a triangular process.[11] She describes the relationship of a "design backward" (from the broadest learning outcome of the institution or discipline to the most specific learning outcomes of an individual lesson) and the "deliver forward" experience (from the students' perspective). Since many institutions are in various stages of developing learning outcomes at each level, librarians have many opportunities to participate in the process of integrating learning outcomes through all levels of the institution.

Assessment
Assessing learning outcomes is an exceptionally complex process and, to guide the process, certain fundamental questions need to be answered: the purpose, the content, and the level of assessment all give direction to the development of assessment measures. A number of models for assessing learning outcomes are being developed and tested on different campuses. Recent literature on higher education

assessment illustrates that some common themes and learning objectives are arising among the general educational outcomes being tested and used. A key resource for monitoring developments in this rapidly evolving area is *Assessment Update*, a journal edited by Trudy W. Banta, Vice Chancellor for Planning and Institutional Improvement at Indiana University-Purdue University at Indianapolis.[12]

In order to assess how library programs and services contribute to student learning outcomes at the student, department, or institutional level, librarians need to inform themselves on how learning outcomes assessment is practiced on their own campus. In addition, the contributions of the library to learning outcomes could be described in a broader context of the students' experiences with the library; not just those limited to the instruction activity.

> *Recent literature on higher education assessment illustrates that some common themes and learning objectives are arising among the general educational outcomes being tested and used.*

Since the library has multiple roles such as service, support for research, and instruction, these roles may need to be addressed separately and by different means of assessment. For example, the LibQUAL+™ project addresses assessment of the quality of service provided by a library while the Assessing ILL/DD project, as well as the MINES methodology proposed to document usage patterns of networked electronic services, both contribute to an assessment of library support for research.

In the movement toward learner-centered education, the emphasis on the student and on learning outcomes is also growing in importance as part of the evaluation of academic programs. Evidence of the contribution of the library to student learning outcomes is therefore an important aspect of demonstrating the value of the library to academic programs and the institution.

Measures
Developing measures to document how the library contributes to the learning outcomes at the department or institutional level has been challenging. Focusing on the course level, a project by ACRL resulted in a report on how librarians in 23 institutions are working with faculty in assessing information literacy learning outcomes.[13] In a study to evaluate an information literacy program, librarians who worked with faculty in the English department noted that the collaboration with the English faculty increased awareness that "information literacy standards are remarkably similar to the standards of critical thinking they try to impart in freshman English."[14] This recognition of the relationship of critical thinking and information literacy

provides a foundation upon which to build mutual learning outcomes.

Measures can be categorized as direct or indirect and within each of these groups, the measures can be quantitative or qualitative. Indirect measures may not be adequate alone, but may enrich or illuminate aspects of what direct measures say about levels of academic achievement.

One indirect measure that has been used by institutions is the College Student Experiences Questionnaire (CSEQ). This survey asks students to indicate the amount of time they spend engaged in various activities that affect learning. It includes eight questions regarding library use. Analysis of the responses shows that "focused library activities" have a "significant" impact on self-reported development of critical thinking.[15] Others found that "students who perceive that their campus emphasizes information literacy gain more in this area, net of other influences."[16] The same researchers also found that the "library experiences of undergraduates are related to select educational purposeful activities" and that the library "appears to be a positive learning environment." However, the results also indicated "the library use does not appear to directly contribute to gains in information literacy and other desirable outcomes." The authors discuss the following possible explanations: CSEQ may not be the best measure for information literacy, "lack of baseline measures" and "a variety of experiences during college…contribute to gains and satisfaction, not just one type of experience."

At the system level, the California State University has launched a three-phase assessment of information literacy skills of the students at all 23 campuses. This project uses both qualitative and quantitative measures and employs multiple methods of assessment. One part of the inquiry includes questions related to information scenarios as well as sets of questions scaled to areas related to information competence "such as library use, research process skills, achievement, presence of reference materials in the home, and computer and media literacy." The "information scenarios were constructed to elicit evidence of information competence when applied to general knowledge, rather than discipline-based knowledge." Responses are measured in terms of "Breadth" (number of different ideas) and "Depth" (number of discrete ideas). In a preliminary report, the author finds that the students that self-reported difficulty finding and using information tended to have lower Breadth and Depth

> *In order to assess how library programs and services contribute to student learning outcomes at the student, department, or institutional level, librarians need to inform themselves on how learning outcomes assessment is practiced on their own campus.*

scores. The study also found that "the greater the usage of the library resources, the greater the Breadth and Depth of the student's responses to the scenarios."[17]

The SAILS Project, at Kent State University and ARL, seeks to develop a standardized test "to measure information literacy skills, gather national data, provide norms, and compare information literacy measures with other indicators of student achievement." The instrument, based on the ACRL standards, is "for programmatic level assessment of information literacy skills that is valid and thus credible to university administrators and other academic personnel."[18] It is possible that the SAILS instrument could be used as a direct measure of information skills development. It would then be necessary to link this to the learning outcomes at the department, program, or institution level.[19]

Performance Indicators

Lindauer identifies several performance indicators that might be used as measures. For example,

- "Perceptions of recent graduates about their information literacy skills training/experience from undergraduate study contributes to their success in graduate/professional programs."
- "Success in applying information literacy skills on the job as perceived by alumni and employers."

The same author suggests that "libraries could strengthen their evidence by triangulating data and supporting information. They could group such measures as selective qualitative descriptions of resources and services with usage data for specific services and with user satisfaction and perceptions of benefit from survey findings. All this could be grouped to show the impact of library and information resources and services…"[20]

Conclusion and Next Steps

The Learning Outcomes Working Group white paper brings together more detail about these and other resources on the state of assessment of learning outcomes especially as they relate to libraries. The goal is that this environmental scan of available research and experience will enable the working group to focus on identifying the most promising mechanisms or instruments that would be useful in demonstrating the research library's direct or indirect impact on learning outcomes at the department or institutional level. A draft of the white paper will be available on the ARL Web site by the end of the calendar year. For more information, contact Sheila Young <sheila.young@asu.edu> or Julia Blixrud <jblix@arl.org>.

—Copyright 2003 Sheila Young

[1] Kenneth R. Smith, "New Roles and Responsibilities for the University Library: Advancing Student Learning through Outcomes Assessment," <http://www.arl.org/stats/newmeas/outcomes/HEOSmith.pdf>.

[2] Mark Battersby and the Learning Outcomes Network, "So, What's a Learning Outcome Anyway?" Vancouver, B.C.: Centre for Curriculum, Transfer, and Technology, July 1999, ERIC Document ED430611.

[3] "Student Learning Outcomes Workshop," *CHEA Chronicle* 5, no. 2 (May 2002), <http://www.chea.org/Chronicle/vol5/no2/Chron-vol5-no2.pdf>.

[4] Battersby, op. cit., p. 8.

[5] Bonnie Gratch Lindauer, "Comparing the Regional Accreditation Standard: Outcomes Assessment and Other Trends," *Journal of Academic Librarianship* 28, no. 1 (2002): 14–25.

[6] <http://www.abet.org/images/Criteria/E1%2003-04%20EAC%20Criteria%2011-15-02.pdf>.

[7] Elizabeth A. Jones et. al., *National Assessment of College Student Learning: Identifying College Graduates' Essential Skills in Writing, Speech and Listening, and Critical Thinking: Final Project Report* (Washington, D.C.: National Center for Education Statistics, U.S. Department of Education, Office of Educational Research and Improvement, 1995).

[8] <http://www.ala.org/acrl/ilcomstan.html>.

[9] <http://www.ala.org/acrl/toolkit/standards.html>.

[10] <http://www.ala.org/Content/NavigationMenu/ACRL/Standards_and_Guidelines/Characteristics_of_Programs_of_Information_Literacy_that_Illustrate_Best_Practices__A_Guideline.htm>.

[11] Mary E. Huba and Jann E. Freed, *Learner-Centered Assessment on College Campuses: Shifting the Focus from Teaching to Learning* (Boston: Allyn & Bacon, 2000).

[12] *Assessment Update*, Jossey Bass.

[13] Gabriela Sonntag, "From Expectations to Results: Library/Faculty Partnerships for Assessing Student Learning Outcomes," presented at an American Library Association Pre-Conference Workshop, Atlanta, June 14, 2002.

[14] Mark Emmons and Wanda Martin, "Engaging Conversation: Evaluating the Contribution of Library Instruction to the Quality of Student Research," *College & Research Libraries* 63, no. 6 (2002): 545–560.

[15] E. Whitmire, "Development of Critical Thinking Skills: An Analysis of Academic Library Experiences and Other Measures," *College & Research Libraries* 59 (1998): 266–273.

[16] George D. Kuh and Robert M. Gonyea, "The Role of the Academic Library in Promoting Student Engagement in Learning," *College & Research Libraries* 64, no. 4 (2003): 256–282.

[17] Kathleen Dunn, "Assessing Information Literacy Skills in the California State University: A Progress Report," *Journal of Academic Librarianship* 28, no. 1 (2002): 26–35.

[18] <http://www.library.kent.edu/sails/>.

[19] The authors have recently published a description of the development of the SAILS instrument. See Lisa G. O'Connor, Carolyn J. Radcliff, and Julie A. Gedeon, "Applying Systems Design and Item Response Theory to the Problem of Measuring Information Literacy Skills," *College & Research Libraries* 63, no. 6 (2002): 528–543.

[20] Lindauer, op. cit.

EVALUATING THE EDUCATIONAL IMPACT OF DIGITAL LIBRARIES

by Martha Kyrillidou, Senior Program Officer for Statistics and Measurement, ARL

The National Science Digital Library (NSDL) Educational Impact and Evaluation (EI&E) Standing Committee is developing a strategy for evaluating the educational impact of digital libraries. As in all new measures work, the lack of common definitions or understood vocabulary is a challenge. Conceptually, however, assessing educational impact is trying to answer the question, "What difference does it make?" in terms of outcomes achieved (e.g., changes in behavior, attitudes, skills, knowledge). Questions of impact are difficult to measure and they usually require extensive and long-term assessment efforts.

What is NSDL?

NSDL is a National Science Foundation (NSF) program with over 100 projects.[1] It is a network of digital libraries and services that are targeting different audiences. A large component of this network is the core integration effort that attempts to create shared technologies, protocols, a central repository, and a core portal (NSDL.org). The mission of NSDL is (a) to encourage and sustain continual improvements in the quality of science, technology, engineering and mathematics (STEM) education for all students, and serve as a resource for lifelong learning and (b) a collective effort to build a national digital library of high quality STEM educational materials for students and teachers at all levels, in both formal and informal settings. ARL and Texas A&M are partners in one of the NSDL projects funded by NSF to explore the development of an evaluation tool for NSDL. (See the article on the ARL e-QUAL project on page 8 of this issue.)

Evaluation in the Context of NSDL

Evaluation of NSDL has been primarily formative in nature with different projects and participants bringing forward their own assessment activities. In addition, there is continuous NSF program evaluation and encouragement to broaden participation of other NSF programs that support the development of assessment capacity.

The emphasis of the EI&E Committee is to develop a community for evaluating digital libraries at a variety of levels and using diverse models. To advance this process a workshop on "Developing an Evaluation Strategy for the Educational Impact of the National Science Digital Library" was held in Washington, D.C., on October 2–3, 2003. ARL's e-QUAL project was represented and a report on EI&E progress will be developed for the ARL community.

[1] Lee Zia, "The NSF National Science, Technology, Engineering, and Mathematics Education Digital Library (NSDL) Program: New Projects and a Progress Report" *D-Lib Magazine* 7, no. 11 (November 2001), <http://www.dlib.org/dlib/november01/zia/11zia.html>.

PROJECT SAILS: STANDARDIZED ASSESSMENT OF INFORMATION LITERACY SKILLS

by Julia C. Blixrud, Assistant Executive Director, External Relations, ARL

Information literacy is the ability to "recognize when information is needed and…locate, evaluate, and use effectively the needed information."[1] This skill set is developed through information literacy programs operated by libraries primarily, but not exclusively, for undergraduates in academic institutions. Are these programs successful? Do they advance the skills of students? How can we measure these programs' success? Kent State University and ARL have partnered on Project SAILS (Standardized Assessment of Information Literacy Skills), which is developing an instrument for programmatic level assessment of information literacy skills that is valid and thus credible to university administrators and other academic personnel.

> *… if university administrators were to be persuaded to allocate resources necessary to grow the information literacy program, librarians must be prepared to substantiate the claim that these skills indeed make a difference institutionally.*

Several years ago, librarians at Kent State University Libraries & Media Services realized that, if university administrators were to be persuaded to allocate resources necessary to grow the information literacy program, librarians must be prepared to substantiate the claim that these skills indeed make a difference institutionally. They must be able to answer such questions as: What are students' entry skills upon admission to the university, and is there a significant change in skill levels from the freshman year to graduation? If there are significant changes in students' information literacy skills, do those skills then have any correlation to student's academic success and retention? The librarians conducted a thorough search of the library literature and determined that the profession was not yet in a position to agree upon the best method for assessing information literacy skills, let alone assert that those skills make a difference.

Librarians need a tool to measure information literacy that could be standardized, is proven to be valid and reliable, contain items not specific to a particular institution or library but assesses at the institutional level, can be administered easily, and provides for both external and internal benchmarking. With such a tool, a library could measure information literacy skills, gather national data, provide norms, and compare information literacy measures with other indicators of student achievement. Libraries would be able to document information literacy skill levels, establish internal and peer benchmarks of performance, pinpoint areas for improvement, identify and justify resource needs, and assess and demonstrate effect of changes in their instructional programs.

Using the Wisconsin Ohio Reference Evaluation Project (WOREP) as inspiration, Lisa O'Connor, Instructional Services Coordinator; Carolyn Radcliff, Head of Reference Services; and Julie Gedeon, Manager of Academic Technology Services Evaluation at Kent State University Libraries & Media Services began using the process of systematic instructional design to develop such a tool in 1998. They chose to use item response theory (IRT) as the measurement model on which to create a new instrument.[2]

A set of questions of varying difficulty level is being produced to measure information literacy. Each question addresses one of the model learning outcomes as identified by the *ACRL Information Literacy Competency Standards for Higher Education*.[3] Use of these outcomes will maximize the applicability of the instrument to a wide variety of academic institutions for internal and external benchmarking. The questions are tested with individual students, in small groups, and in field trials (which closely emulate actual survey administrations).

While the SAILS survey is intended to be delivered in electronic form over the Web, it can be delivered using paper and pencil. Using the Web, a random set of about 45 questions drawn from the instrument's data bank of now approximately 150 items are delivered at the time the test is administered to a student. Project participants can choose the most appropriate testing population for their institution as long as the students are undergraduates.

Institutions can test students in one of three situations:

- Highly monitored: The test-taking is proctored and the procedures are consistent across all students. Students check in and out of a specified testing site and the proctor is aware of the testing. A variation of this is when students take the test as a group during class time.
- Loosely monitored: Students come to a specified testing site, but do not check in or out.
- Unmonitored: Students take the test anytime and from anywhere.

The answers are collected and analyzed using WINSTEPS, a modeling program created by researchers at the Mesa Institute at the University of Chicago.

The Kent State University staff received a grant in fall 2002 from the Institute of Museum and Library Services (IMLS) for Project SAILS. The three-year grant is supporting continued development of the instrument and testing at other institutions. It will also enable project staff to create the foundation of a results database that will address internal and external benchmarking. The grant also includes funding for the commissioning of three discipline-specific assessment modules, which will serve as a model for future development of other discipline-specific modules.

Since there is significant overlap between student learning outcomes and information literacy, ARL's Learning Outcomes Working Group recommended that the SAILS Project become one of the projects under the ARL New Measures Initiative. The ARL Statistics and Measurement Committee endorsed SAILS and positioned it as an activity within the Learning Outcomes Working Group agenda. As with other New Measures projects, participating institutions pay fees to cover some of the costs of tool development and data reporting.

Project SAILS is a collaboration between the Information Services and the Libraries & Media Services divisions at Kent State University. Mary Thompson, Interim Liaison Librarian for Business, serves as Project Coordinator; Julie Gedeon manages data measurement and statistics; Lisa O'Connor is in charge of test development; Carolyn Radcliff handles test administration; and Rick Wiggins is the Web Programmer. In addition to project participants, an advisory council provides input for the survey items.

ARL is responsible for coordinating and managing participation for all institutions whether they are ARL members or not. This includes calls for participation; making arrangements for planning meetings, training workshops, and follow-up meetings; and marketing and public relations for the project. Julia Blixrud is the ARL liaison for Project SAILS.

The IMLS-funded activities are divided into three phases. Each phase is examining both the survey items as well as test administration. Participants are also closely involved in working with Project SAILS staff to develop data reports. Phase I was conducted in spring 2003 at seven institutions: University of Arizona, Auburn University, Oregon State University, San Jose State University, Virginia Polytechnic Institute and State University, Washington State University, and Weber State University. Representatives of those institutions met in June 2003 to review the results of data analysis and to share experiences regarding testing situations.

Phase II covers the fall 2003 and spring 2004 semesters and over 40 institutions have agreed to participate.

Phase III is scheduled for fall 2004 and spring 2005, with a projected participation of 100 libraries. A call for participation for Phase III will be issued in January 2004.

The rapidity by which institutions signed up for Project SAILS and the enthusiasm for participation signals that libraries are looking for tools that can help them determine the effectiveness of their information literacy programs. It is another example of how the leadership of one library working collaboratively with other interested libraries can create a useful, standardized tool to benefit the larger library community.

More information on Project SAILS can be found at <http://www.projectsails.org/>.

[1] American Library Association, Presidential Committee on Information Literacy, "Final Report" (Chicago: American Library Association, 1989), <http://www.ala.org/Content/NavigationMenu/ACRL/Publications/White_Papers_and_Reports/Presidential_Committee_on_Information_Literacy.htm>.

[2] Lisa G. O'Connor, Carolyn J. Radcliff, and Julie A. Gedeon, "Assessing Information Literacy Skills: Developing a Standardized Instrument for Institutional and Longitudinal Measurement," in *Crossing the Divide: Proceedings of the Tenth National Conference of the Association of College and Research Libraries* (Chicago: ACRL, 2001): 163–174, <http://www.ala.org/Content/NavigationMenu/ACRL/Events_and_Conferences/oconnor.pdf>; and Lisa G. O'Connor, Carolyn J. Radcliff, and Julie A. Gedeon, "Applying Systems Design and Item Response Theory to the Problem of Measuring Information Literacy Skills," *College & Research Libraries* 62, no. 6 (Nov. 2002): 528–543.

[3] <http://www.ala.org/Content/NavigationMenu/ACRL/Standards_and_Guidelines/Information_Literacy_Competency_Standards_for_Higher_Education.htm>.

LIBRARY IMPACT ON RESEARCH

DOCUMENTING USAGE PATTERNS OF NETWORKED ELECTRONIC SERVICES

by Brinley Franklin, Director of Libraries, University of Connecticut, and Terry Plum, Assistant Professor, Simmons Graduate School of Library and Information Sciences

A new initiative now being considered by the ARL Statistics & Measurement Committee is a Web-based survey that a library could administer to collect reliable data on the usage of networked electronic services from both within the library and remotely through the Web. The idea is to build on the experience of studies conducted at five academic health science libraries and two main libraries between 1999 and 2003. In those studies, more than 15,000 library users were surveyed as they accessed their libraries' networked electronic services. The survey documented *why* patrons were using electronic resources in those libraries and tracked the differences between in-house and Web usage, comparing the *location* of the user and their *status* (graduate student, faculty, undergraduate, etc.) with the *purpose of the use*.

The methodology measured both in-house and remote usage in a way that allowed the costs of providing networked electronic services to be distributed among the university's primary functions. In this study, those functions were defined in four categories: sponsored (funded) research, instruction/education/departmental research, patient care, and all other activities. The survey is a pop-up screen administered randomly to users of networked electronic services throughout the year using a statistically valid sampling plan.

The study, now called "Measuring the Impact of Networked Electronic Services" (or MINES), was designed and undertaken by the authors. The researchers concluded that although the computing environments and survey implementations in each library were different, the methodology and overall approach was workable at each library. As a result, they are confident that the data collected are somewhat comparable, recognizing that there will always be distinguishing characteristics driven by the local user population and institutional characteristics.

Experience also led the researchers to conclude that the most effective means to meet the Web survey criteria was to run all access for networked electronic resources through a gateway that authenticated access

FIRST PAGE OF SURVEY

Networked Electronic Services Library User Survey

This survey is being conducted by the University in order to assess usage of the Library's electronic services.

All responses are confidential.

After completing the survey, you will be connected to the service you selected.

Thank you for your help.

Patron Status: [Select Patron Status ▼]

Affiliation: [Select Affiliation ▼]

Location: [Select Location ▼]

Purpose for Using Online Resources:

- ◉ A. Sponsored (Funded) Research —Definition
- ◉ B. Instruction/Education/ Departmental Research —Definition
- ◉ C. Patient Care —Definition
- ◉ D. All Other Activities —Definition

Please fill out Status, Affiliation, Location, and Purpose Fields.

[Submit Response]

and passed on the request. Such an arrangement prevents lost data due to bookmarks, non-library Web pages, and other non-library routes to access library networked electronic services.

Initial use of the methodology produced a number of findings useful for guiding decisions on how to manage networked electronic resources. For example, results from the first four academic health science libraries showed that:

- There were approximately four remote networked electronic users for every in-house user.

- Remote users were demographically different from in-library users; the 4-1 ratio was even higher for the category of users who identified themselves as faculty/staff/research fellows.

- The purpose of use reported for remote usage of networked electronic resources was significantly different from in-house usage. For example, those using networked electronic resources to conduct sponsored research do so most intensively from on-campus but not from in the library. Those using these services for instruction/education/non-sponsored research were more likely to be in the library than people using the resources for other purposes.

- Patrons using electronic services for sponsored research represented 34% of the usage in the health science libraries, as compared to 16% in the two main libraries.

The study also found that, despite proxy servers, modem pools, and similar remote enabling services, the patrons being surveyed went to the university to use their library's networked services. That is, when they used these services they were either on campus or in the library many more times than off-campus. However, the users' purposes for using electronic resources ranged significantly among the four libraries, underscoring the need to consider local conditions as libraries have different characteristics and user populations.

The authors concluded that MINES's Web-based survey methodology employed at the seven libraries could serve as a model for similar user studies at other libraries. The discussions now underway within the ARL Statistics and Measurement Committee focus on how to take a next step to apply this methodology to the networked electronic services offered by a more complete complement of research libraries.

A full description of the research project and its findings was first delivered at the Northumbria Lite conference, part of IFLA in Glasgow in 2002. See Brinley Franklin and Terry Plum, "Networked Electronic Services Usage Patterns at Four Academic Health Sciences Libraries," *Performance Measurement and Metrics* 3, no. 3 (2002): 123–133, <http://www.arl.org/stats/newmeas/emetrics/Franklin_081102.pdf>. Subsequent findings were recently presented by the authors in September 2003 at the Conference on Users in the Electronic Information Environment in Espoo, Finland.

COST EFFECTIVENESS OF LIBRARY OPERATIONS AND SERVICES

ASSESSING ILL/DD SERVICES STUDY: INITIAL OBSERVATIONS

by Mary E. Jackson, Director of Collections and Access Programs, ARL

What are the characteristics of a cost-effective interlibrary loan (ILL) and document delivery (DD) operation? To get a clearer answer to this question, ARL has undertaken a third study of the performance of ILL/DD operations in North American libraries. The current study, Assessing ILL/DD Services, collected data on the 2001–02 performance of mediated and user-initiated ILL/DD operations in 72 research, academic, and special libraries.

New to this study is a focus on user-initiated, or unmediated, ILL services. User-initiated ILL is defined, for this study, as requests that are initiated without ILL staff involvement and received directly by a potential supplier. This is the first comprehensive study of a range of user-initiated services: INNReach, URSA, Loansome Doc, ILLINET Online, RAPID, and user-initiated use of commercial document delivery suppliers. In addition, the project collected data on the performance of local document delivery services in 25 libraries.

The project measured unit cost, fill rate, and turnaround time for mediated and for user-initiated interlibrary borrowing and lending. Cost data includes the costs for staff, network/communications, delivery, photocopying, supplies, equipment, and borrowing fees. Turnaround time for borrowing, as in the previous study, was measured in calendar days from the date the user submitted the request to the date ILL staff notified the user that the item was available for pickup (or mailed the item to the user). Turnaround time for lending, a measure not included in the previous study, captures the elapsed time between a library receiving and shipping a request.

When the cost data are combined with performance data on turnaround time and fill rate, it is possible to identify the characteristics of the most cost effective interlibrary borrowing and lending operations and to showcase good practices that could benefit other libraries.

Results from Previous ILL/DD Studies

In 1992, ARL and the Research Libraries Group (RLG) collaborated in a joint project to collect detailed

ASSESSING ILL/DD SERVICES STUDY
RESULTS FOR ARL LIBRARY PARTICIPANTS, MEAN VALUES

	Borrowing		Lending	
	2002	**1996**	**2002**	**1996**
Unit cost, Mediated	$17.50	$18.35	$9.28	$9.48
Unit cost, User-initiated	$2.04–$14.70	NA	$3.47–$12.06	NA
Fill rate, Mediated	86%	85%	58%	58%
Fill rate, User-initiated	84–91%	NA	76%–88%	NA
Turnaround time, Mediated	9.29 days (loans) & 6.23 days (copies)	16.9 days (loans) & 14.9 days (copies)	1.75 days (loans) & 1.37 days (copies)	NA
Turnaround time, User-initiated (loans)	2.30–8.23 days	NA	0.57–1.45 days	NA

information on 1991 costs incurred by research libraries for ILL transactions.[1] That study found that research libraries spent an average borrowing unit cost of $18.62 and lending unit cost of $10.93.

In 1995, ARL conducted a second study that built on the 1992 study by measuring not only the cost but also several performance attributes of ILL/DD operations in research libraries and, for comparison purposes, in college libraries.[2] This study found that research libraries spent an average of $18.35 on a borrowing transaction, $9.48 on a lending transaction, and took an average 15.6 calendar days to receive a borrowing request.

The Current Results
The accompanying chart presents the current findings for the 59 participating ARL libraries and compares the results with those for the 97 ARL participants in the 1996 study. It is important to remember that different libraries participated in both studies, which may result in discrepancies when comparing the two different studies. Additional analysis will be undertaken to compare the results of the 46 libraries that participated in both the 1996 and 2002 studies and the 29 libraries that participated in all three ARL ILL/DD studies.

The current study confirms informal, institution-specific studies and speculations that user-initiated services provide better service to users than mediated ILL. Overall, user-initiated services have lower unit

costs, higher fill rates, and faster turnaround times than mediated ILL.

Mediated borrowing has improved compared to the 1996 study. The current study reports a unit cost for mediated borrowing of $17.50 in 2002 dollars. When adjusted for inflation, that unit cost would represent $14.88 in 1996 dollars, compared with $18.35 in the previous study and represents a 19% reduction in constant dollars. Staff costs now represent less than 60% of the unit cost. Photocopy as a percentage of the overall lending unit cost has decreased and borrowing fees now represent a larger percentage of the borrowing unit cost.

Turnaround time for mediated borrowing has decreased as has the volume of mediated borrowing and lending compared with the 1996 study.

Bruce Kingma, Associate Dean, School of Information Studies, Syracuse University, and Tom Delaney, Head of ILL at Colorado State University, are collaborating with Mary Jackson, ARL, in the conduct and analysis of this study. For more information, contact Mary Jackson <mary@arl.org>.

[1] *ARL/RLG Interlibrary Loan Cost Study: A Joint Effort by the Association of Research Libraries and the Research Libraries Group* (Washington, D.C.: ARL, 1993).

[2] *Measuring the Performance of Interlibrary Loan Operations in North American Research & College Libraries* (Washington, D.C.: ARL, 1998).

TIME AND COST ANALYSIS: STAFF ALLOCATIONS PROJECT

by Dilys Morris, President, Chronos Group

In February 2004, Notre Dame and Vanderbilt Universities will begin studying how staff time and costs are allocated across the library. The two libraries have worked together with the Chronos Group to identify library-wide time centers. They will use the Chronos Group's TCA DecisionBase methodology and software to collect and analyze data. Vanderbilt will test the time centers during the fall of 2003.

The goal of the study is to understand time and costs at the macro, or center, level. Time data will also be collected at a micro, or task, level. Both libraries will have a mix of macro and micro data since some administrative areas of each library prefer to collect data at the micro level. Since the centers represent library functions, it will be possible to see how time is distributed by function across the libraries. Additionally, because reports can be sorted by any administrative unit, the study will show all functions performed by a given unit, including those that fall outside the normal work of that unit. The libraries will be able to examine how the actual time spent compares with their priorities and expectations, and evaluate whether it is possible to allocate more time to higher priorities. Finally, it will be possible to compare time allocations (both actual time and proportions) with those of a similar institution.

The libraries will code staff according to three broad employment groups: exempt from the 40-hour workweek (e.g., professionals), nonexempt, and students. Comparisons by these employment groups will be possible and each library will enter finer classification types for more specific local reporting.

The TCA DecisionBase time-cost benefits analysis methodology combines a longitudinal approach and time sampling. Four sample weeks will be selected annually. For the remainder of FY 2004, two weeks only will be sampled: February 23–29 and May 10–16. To facilitate comparisons between the libraries, no holiday weeks are included in the sample.

During a sample week, all staff time will be tracked by the time centers. The centers are divided into two groups: *Product/Service Centers* and *Overhead Centers* (see below for a list of centers). An *Excluded Center* tracks time that is not used in reports, such as leave without pay.

The separation of Overhead Centers from Product/Service Centers will allow the libraries to get a clear picture of their overhead costs. Reports identify the overhead costs for each layer in the organization. Costs of each administrative group are spread across all units supervised.

Costs for a given center consist of employee salary and fringe benefits multiplied by the time spent in that center.

LIBRARY-WIDE TIME CENTERS

Product/Service Centers
- Acquisitions
- Catalog Maintenance
- Cataloging
- Circulating Materials
- Collection Development & Management
- Conversion or Major Projects
- Development & PR
- Digital Collections
- ILL & Document Delivery
- Instruction
- Preservation
- Storage & Stacks Maintenance
- User Assistance
- Volume Preparation

Overhead Centers
- Administrative & Support Services
- Automation & Systems
- Paid Leave

Excluded Center
- Time Not Counted

Employment data for each employee is entered once, but can be updated for subsequent sample weeks as needed. Each employee has a Position Code that reflects the organizational structure. Individuals who work in multiple organizational units are assigned multiple Position Codes.

To collect the time sampling data, every employee will manually complete a time sheet that identifies the time by centers for the entire week. If tasks within centers are used, the time spent at each task will be recorded. Time is recorded in 15-minute increments, and rounded to the nearest quarter of an hour. Break time is not recorded; instead, employees will spread non-working time over the centers or tasks worked.

TCA DecisionBase calculates the hourly salary of each employee for every sample week. The hourly salary is multiplied by task or center time to arrive at costs for each employee. The software sums time and costs. Output measures can be entered for each sample week and are used to determine unit costs and times per task or center. These measures (e.g., titles cataloged, materials circulated, questions answered, etc.) can be entered for any task or center.

Reports give both time and costs for centers and tasks, and can be generated for any defined administrative unit. Additionally, reports can be generated for the three employee groups mentioned above. Weekly and projected annual reports are produced.

—Copyright 2003 Dilys Morris

For further information, please contact Dilys Morris at the Chronos Group <dmorris@iastate.edu>, Flo Wilson at Vanderbilt <flo.wilson@vanderbilt.edu>, or Jo Bessler at Notre Dame <Joanne.M.Bessler.1@nd.edu>.

LIBRARY FACILITIES AND SPACE

EVALUATING THE UTILIZATION OF FACILITIES IN ARL LIBRARIES

by Clifford H. Haka, Director of Libraries, and Jim Hensley, Facilities Manager, Michigan State University Libraries

How can libraries tell if they are effectively utilizing their space? Is there a way to measure how we currently utilize the physical space in research libraries? collect data and compare it across libraries? and then assess the impact of various space allocation decisions on library activity such as building use and circulation?

For example, two libraries might pursue radically different approaches to shelving materials, with one using traditional shelving while the other utilizes compact shelving. This would result in the allocation of space for shelving in one library being much higher than in the other library, and enable us to assess if this has an impact on use, such as the total number of circulations. Similar examples might include the extent to which libraries subscribe to electronic journals, thereby limiting the space required for storage of physical journal titles, or the degree of outsourcing, which would decrease the amount of space required for staff operations.

No single measure, such as the examples suggested above, would allow libraries to respond to the overall question of the future role of physical library facilities. Such measures should, however, provide insight regarding those factors that encourage, or possibly discourage, the utilization and subsequent effectiveness of library facilities.

While a review of space utilization along these lines might not fit the classic definition of a "performance measure," this topic was, in fact, identified as an area of interest during the initial ARL-sponsored New Measures Retreat held in Tucson in 1999. Could procedures be developed that might assist in the assessment of current space utilization within ARL libraries? No "ready-made tool" exists that would enable a library to review this matter objectively and expeditiously. However, many at the retreat felt that it would be interesting, and potentially useful, to conduct a detailed review of space utilization within several ARL libraries to determine if trends, or possibly of more interest, abnormalities, exist in regard to the manner in which ARL libraries use their space.

Library Space Evaluation at Michigan State University

To get this ball rolling, the Michigan State University (MSU) Libraries undertook to establish detailed CAD files for the Main Library facility. This was done by employing CAD 2000 software.

CAD files containing the basic layout of the library building were obtained from the Physical Plant Engineering Department. Unfortunately, these files proved to be incomplete and in some cases inaccurate. Therefore it was necessary to verify dimensions by re-measuring many areas before proceeding. Following the establishment of an accurate interior, additional elements such as stacks, service points, and furniture were added to the drawings. (As an aside, once completed these documents have proven to be extremely useful for space-planning projects, especially when units move to new locations. Precise plans can be developed, reviewed, and modified without the tedious and time-consuming sketching and re-sketching, often inaccurately, that was the previous norm.)

The next step was to define different types of spaces found within the typical library. Two primary resources were used, starting with Metcalf.[1] Using guidelines contained within Metcalf, assignable and non-assignable space was identified throughout the building. The identification of non-assignable space—such as mechanical, custodial, and telecommunication areas—is an important first step as these spaces cannot typically be used, but they still make up a significant percentage of the overall square footage of all buildings. The next step was to discern what functions had been assigned to the remaining space. To help with these definitions, a second resource, published by the Buildings and Equipment Section of LAMA,[2] was utilized, resulting in the following seven categories:

- **Non-Assignable:** Areas designated for other uses—structural elements, HVAC chases, electrical vaults, telecommunication closets, custodial areas, etc.
- **Service:** All areas that are points of contact between staff and patrons, or are areas for patron self-service such as copy machines and telephones.
- **Staff Area:** Office space, work areas, and facilities-support areas such as loading dock and mailroom.
- **Stacks:** All areas designated for storage of books and materials regardless of format. For instance, microfilm cabinets fall within this category.
- **Seating:** All seating without network capability.
- **Networked Seating:** All other seating that has network capability.
- **Traffic:** These areas are used primarily for movement around the building and for no other purpose. Main corridors and hallways to elevators, service areas, and restrooms were included.

It would certainly be possible to divide many if not all of these categories into countless subdivisions. However, every division requires further definition and increases the complexity of compiling the data for a library. Therefore,

in the hope that other ARL libraries might be willing to replicate these efforts to provide comparative data, we strove for a limited number of straightforward and unambiguous categories.

Seating serves to depict this dilemma. Differentiating between carrel seating, table seating, and "casual" seating is not that difficult. When one adds the element of "laptop-ready seating" however, which cuts across all these seating types, it was decided, at least for the purposes of coding the MSU floor plans, to collapse seating types into networked and all other seating. It was subsequently decided to depict additional detail via other techniques. For example, a pie chart was developed to display a more detailed range of seating options, as well as their relative availability (see accompanying chart). To use such detailed distinctions when coding floor plans would, however, yield an indecipherable maze that would be exceptionally daunting and time consuming to maintain with any degree of accuracy. Hence, we pursued fewer categories for the overall coding of the floor plans with detail represented via alternative charts.

It would seem that delineating and marking the plans into these seven categories would be relatively straightforward at this point. This proved not to be the case. This part of the project was more time-consuming than anticipated. For example, defining spaces for seating that are adjacent to walkways is difficult to measure in that the chair push-back space from the table or carrel frequently encroaches into the area defined as a walkway. Specifically the push-back from a 30-inch work surface carrel was defined in the LAMA document as 4 ft. 6 in., a significant amount of floor space that frequently spilled into walkway space. In a very real sense this was the confirmation that original space allowances had been discarded, at least in this instance at the MSU Library. This space was then defined as dual purpose and it became an issue to address. In this particular case it was decided that the hallway space took precedence and was measured as such. However, there were many such measurement decisions that made "fine tuning" of our CAD document somewhat challenging. Listing all of the various definitions and/or conventions we considered too exhaustive for the purposes of this paper. However, most are found in the two sources cited.

Upon completion of floor plan coding,[3] it was relatively straightforward to produce charts and graphs to depict specific aspects of building use. In turn, this enabled the calculation of average space allocations for items such as seating. For example, one calculation made across all types of seating found that an average of 31 square feet per seat has been provided, which falls well within the standard established by Metcalf. Information such as this can prove useful in considering issues such as, "Have too many seats been introduced into too small

MSU SPACE UTILIZATION, SEATING BY TYPE

Seating/special format (microfilm) 0.44%
Seating/computer 10%
Seating/carrel/netlink 9%
Seating/fac/grad/carrel/netlink 3%
Seating/a/v 2%
Seating/group 42%
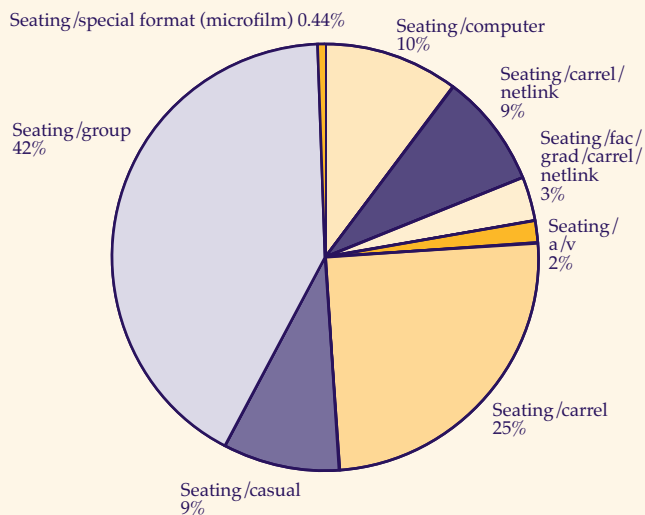Seating/carrel 25%
Seating/casual 9%

Chart created by Jin-hee Kim, MSU Graduate Student

of a space?" Calculations for book-stack capacity would be similarly instructive and potentially useful in argumentation for additional space.

The creation of the CAD drawings was the most complex and time-consuming portion of this project. However, the hours spent on this task have paid huge dividends for space-planning discussions. The process is, of course, ongoing as staff and users move and relocate chairs, seating, and other items.

The MSU Libraries plan to update and review these documents on a regular basis. We hope that we can engage one or more ARL libraries to become partners with us to further define this process. Our goal is to create a tool that will prove helpful for comparative purposes and to assess the impact of various space allocation decisions on library activity. If a group of ARL libraries can be identified to compile comparable data, this could provide "benchmark measures" against which other libraries could compare their situations. We look forward to this possibility.

—Copyright 2003 Clifford H. Haka and Jim Hensley

[1] Keyes D. Metcalf's *Planning Academic and Research Library Buildings* (New York, McGraw-Hill: [1965]) has been employed in countless library construction projects over the past 35 years.

[2] American Library Association, Library Administration & Management Association, Buildings and Equipment Section, Functional Space Requirements Committee, *Building Blocks for Planning Functional Library Space* (Lanham, Md.: Scarecrow Press, 2001).

[3] MSU's coded floor plan and samples charts and graphs produced with the CAD software are available on the ARL Web site <http://www.arl.org/newsltr/230/space.html>.

ORGANIZATIONAL CAPACITY

CREATING A CULTURE OF ASSESSMENT: THE UNIVERSITY OF ARIZONA EXPERIENCE

by Carla Stoffle, Director of Libraries, and Shelley Phipps, Assistant Dean for Team and Organizational Development, University of Arizona Library

Especially given her leadership role in launching the ARL New Measures Initiative, it is not surprising that Carla Stoffle has led the University of Arizona Library to be an early adapter of performance measures. The article below describes how one research library is assessing itself by implementing a performance measurement system and integrating the results of the New Measures Initiative projects as they become available. The benchmarks identified and described by Arizona are made public to provide other libraries with a point of departure for measuring their own operations. This article is reprinted with permission from the authors and the Editor of Library Issues, *where it appeared in the March 2003 issue, vol. 23, no. 4.*

In addition to the national efforts that have set a new direction for performance measures in academic research libraries, individual institutions are experimenting with using assessment techniques to quickly learn about changes in local users' expectations. The importance of implementing an organizational performance measurement system has been one conclusion of such experiments. One institution, the University of Arizona Library, created a Performance Effectiveness Management System (PEMS) in 1998.

PEMS is designed to align individual, unit, and whole organization efforts with strategic goals set by the Library and to provide measures that indicate success, progress, and the need for improvement. Developing a system approach to measurement helps develop an internal culture of assessment where decisions are guided by facts, research and analysis, and where services are planned and delivered to maximize positive outcomes for customers and stakeholders. Staff are encouraged:

- to care about what results they produce;
- to value the actual impact they have on the educational or research process; and
- to know how these results relate to user expectations.

Creating this culture can ensure that there is a commitment to continuing assessment, development of new tools, and use of data for continuous improvement. Without an emphasis on culture change, barriers will remain to the full acceptance of the use of performance measures.

PEMS: How Does It Work?

The PEMS system requires units to continuously assess needs of their assigned faculty and student groups. Using varying techniques, teams engage in discovering what is most important to their customer groups about the services they provide. They use this information to formulate standards or performance targets for the particular service activity. They gather data on progress toward these standards and report periodically to the Library Cabinet and Strategic Planning Team.

After analyzing the needs and assessing where improvement can make a difference, the teams either engage in specific projects to increase their effectiveness or assign individuals the responsibility for increasing the amount or quality of their individual work that results in outcomes for students and faculty. Thus the system supports individual goal setting that furthers the unit's service effectiveness related to what customers have identified as important. Strategic, future focus is structured into the system as the teams start the year's planning by understanding their role in achieving the Library's three to five-year strategic goals.

As teams create their strategic framework for the year, they set outcome, output and quality measures to assess success. In some cases outcomes are measured directly with customers. In other cases, needs assessments lead to the creation of output or quality measures that can be deduced from the identified need. For example, if a team assesses that only a limited portion of its potential user group is using its services, they may set a standard to increase the number of instructional sessions or increase the number of students reached. If they recognize that limited budgets require selection of materials that have a high potential for usage, they may set a standard that measures actual usage of material purchased. If they identify that timeliness is an important aspect of the service they give, they may set a standard for turnaround time, from request to delivery. If they identify that accuracy of information is critical to the users of their service they may set a related standard.

Examples. Examples of outcome measures may include: 95 percent of the serials acquired in science/engineering areas are used at least once each year.

Some standards use increased output as surrogate measures for an intended outcome. For example, 95 percent of UA faculty will receive information about

intellectual property rights and copyright (the intended outcome is that faculty are able to use this information in their teaching and publication efforts).

Some standards relate to the user's need for access or service in a timely fashion or a need for accuracy that will facilitate access: 90 percent of traditional reserve requests at all sites will be available to customers within 48 hours after receiving the request.

Assessment with users has led to process improvement efforts that have decreased cycle time, increased amount and quality of service, and saved hundreds of thousands of dollars that have been reallocated to the purchase and refresh of technology and the implementation of a competitive salary structure.

Learning the tools and methods for assessing performance has been a challenge. The ARL New Measures Initiative has afforded the opportunity to use more reliable measurement methods that will also yield peer benchmarking data. The Library's Strategic Plan now incorporates measures that can be derived from the LibQUAL+™ instrument as well as measures based on team data gathering. For example:

...teams engage in discovering what is most important to their customer groups about the services they provide.

- Degree to which access to electronic information exceeds customers' minimum expectations (from LibQUAL+™): 15 percent increase within 5 years.
- Scholarly electronic resources newly developed by the Library and available remotely to customers (120 new resources within the next 5 years).

Using Data for Strategic Decision Making

Information from the first pilot years of LibQUAL+™ indicated that users desired delivery of electronic information to the desktop and that the Library was not reaching even minimum expectations. A strategic decision was made to create a Document Delivery Team and set standards for electronic delivery of interlibrary-loaned material and reserve articles. Cost information from the ILL/DD project was used to research best practices and to join with other libraries using more efficient technology. This enabled customers to order interlibrary loans directly from other libraries, which reduced turnaround time. A systems analysis project team was formed to research the capability of present systems to provide the infrastructure necessary to improve access in the future. As a result, the UA Library is now participating in the development of the "Scholar's Portal" software with a number of ARL Library partners.

Access to electronic information is indicated as highly desired by LibQUAL+™ respondents and confirmed by team interviews and surveys. This information also contributed to the setting of the standards in the Strategic Long Range Plan and influenced the UA Library's decision to invest a larger portion of its information access budget in electronic information resources.

Evaluation of the University's budgetary situation has led to plans to increase consortial purchases as a way of reducing overall costs to the University and a standard was set to save at least $100,000 per year on such purchases. The UA Library will be joining the SAILS project to research more efficient ways of measuring learning outcomes. As the Library builds its new technological infrastructure, we will be using the research from the E-Metrics project to assess the effectiveness of our electronic access systems—from the customer point of view and in comparison with peers involved in this national effort.

Fulfilling our Educational Mission

The PEMS system and its ability to use the results of national initiatives such as the "New Measures Initiatives," enables the Library to demonstrate to the campus that all resources and staff efforts will be focused on the changing expectations of our users by providing value-added services. Incorporating efficiency measures and developing process improvement projects also demonstrates the Library's commitment to maximizing return on investment. Accountability is demonstrated at the institution, unit and individual level.

Use of performance measurement is a way of ensuring that academic libraries retain their ability to perform their special mission within the educational process. These include:

- Valuing freedom of access to information,
- Eliminating censorship,
- Providing equitable access to all levels of users, and
- Increasing the information literacy of the students and faculty in this increasingly complex scholarly communication system.

Developing systems and using evolving tools created by ARL-sponsored projects and others, increases the ability to measure and communicate performance outcomes and the effect of libraries on the quality of teaching and learning processes in our institutions.

—Copyright 2003 Mountainside Publishing Co., Inc.

USING DATA TO MAKE CHOICES: THE BALANCED SCORECARD AT THE UNIVERSITY OF VIRGINIA LIBRARY

by Jim Self, Director of Management Information Services, University of Virginia Library

In early 2001 the University of Virginia (U.Va.) Library began implementation of the Balanced Scorecard. This initiative continues a long history of collecting and using data to improve services, collections, and processes at the U.Va. Library.

The Balanced Scorecard (BSC) allows an organization to focus on a relatively small number of carefully chosen measurements. These measures are aligned with the library mission and strategies, and they provide a quick but comprehensive picture of organizational performance.

The concept dates from a 1992 article in the *Harvard Business Review* by Kaplan and Norton.[1] The scorecard was designed for private business, but is increasingly being used by governmental and nonprofit organizations. In 2001 ARL and OCLC sponsored a three-day forum on performance measures, which featured the Balanced Scorecard.[2] Several German libraries, under the leadership of the German Research Council, have implemented the Balanced Scorecard.[3] However, implementation in North American libraries has been slow to develop.

The Balanced Scorecard examines the organization from four perspectives: User, Finance, Internal Processes, and Learning and the Future. Four to eight measurements, or metrics, are devised for each category or perspective. Each metric has a specific and unambiguous target or set of targets. For example, two of the metrics adopted by U.Va. for 2002–03 are described below. Each of these metrics contains two targets. Target1 indicates complete success in achieving the metric; Target2 indicates partial success.

To recruit, develop, and retain productive, highly qualified staff, the library established a metric for the *Retention Rate of Commended Employees*:

Target1: 95% of commended employees remain on the staff.

Target2: 90% of commended employees remain on the staff.

Method: Staff members who have received at least one extraordinary contributor write-up and faculty who have received at least one "extra-meritorious" or "highest" salary increase over the past three performance cycles are identified at the beginning of each fiscal year. The library may also take other forms of recognition into consideration. At the end of the year, we will determine what percentage of these employees still work at the library.

To provide resources and services that have a high ratio of value to cost, the library has established a metric for *Expenditures for Circulating Monographs as a Proportion of All Monographic Expenditures:*

Target1: Monographs that circulate within the first two years should account for at least 60% of the monographic budget.

Target2: The circulated items should account for at least 50% of the monographic budget.

Method: A yearly study similar to the 1993–95 baseline study will be conducted. The study will track purchased monographs for two years to determine if a customer uses them.

The U.Va. Library is using this technique for several reasons. In essence the Balanced Scorecard enables us to gain better control of our statistical operations. By limiting the number of scorecard metrics, it forces us to decide what is important, and to identify those numbers that truly make a difference. It also introduces some balance into our statistical work. Like many libraries, we have collected many data regarding resources and user services, but other areas have not received the same attention. The BSC compels us to take a focused look at finance, internal processes, and the future. Another important aspect of the BSC is the assigning of targets or goals. We not only decide what measures are important; we also state what constitutes success for each measurement. A final rationale for the scorecard is its intelligibility; each year we can literally see how well we have done. We get a clear picture of organizational performance in the form of a set of pie charts. We have one pie chart for the overall results, and individual charts for each of the four perspectives. We can evaluate, at a glance, the performance of the library. It is equally easy to compare the four perspectives and to note where we are doing well, and where we need to improve.

Choosing the metrics may be the most crucial part of the process. Once a metric is established, it has organizational approval and recognition—as the saying goes: what gets measured, gets managed. The choice of metrics depends upon the values of our organization. There must be some agreement as to what matters, what counts within the library. Adopting the Balanced Scorecard requires the library to confront issues in a more rigorous fashion, to make some hard decisions, and to make those decisions explicit.

The subject matter of the metric (e.g., reference service, cataloging, fundraising, interlibrary loan) is obviously important, but not the only issue. The nature of the measurement is equally important. Do we count the number of times we perform a task? Do we measure the cost of performing the task? Do we calculate the

time it typically takes to do the task? Or do we survey our customers and ask them how well we performed the task? Any of these techniques might be appropriate, but our choice of techniques is reflective of the library's priorities.

The metrics also reflect practical considerations. Collecting data generates costs, and we attempt to minimize those costs. We do not want the Balanced Scorecard to be a burden; it should not appreciably increase our costs—either in terms of labor or monetary expenditures. Whenever possible, we use existing measurements, incorporating them into the process as scorecard metrics. The second choice is to use data that can be mined efficiently from databases. The third choice is collection through sampling—preferably carried out by student employees. Only as a last resort, do we ask operational ("front line") staff to spend their own time collecting data.

U.Va. has collected scorecard data for two fiscal years. We are currently completing the FY03 tally, and will soon communicate the results to library managers and staff. At the same time we are collecting data for FY04, and starting the planning for FY05.

The process is still evolving, but thus far we are very pleased with the Balanced Scorecard as a management tool. It has improved our statistical focus, clarified and balanced our organizational values, and helped us develop a culture of assessment. We think its utility will only increase as we become more experienced in its implementation.

—Copyright 2003 Jim Self

Additional information can be found at <http://www.lib. virginia.edu/bsc/> or by writing to <self@virginia.edu>.

[1] R.S. Kaplan and D.P. Norton, "The Balanced Scorecard—Measures that Drive Performance," *Harvard Business Review* 70, no. 1 (1992): 71–79.

[2] ARL/OCLC Strategic Issues Forum, "Toward Rethinking Academic Library Performance in the Digital Age," Tempe, Arizona, February 16–18, 2001, <http://www.arl.org/olms/forum2001.html>.

[3] R. Poll, "Performance, Processes, and Costs: Managing Service Quality with the Balanced Scorecard," *Library Trends* 49 (2001): 709–717.

DATA-INFORMED DECISION MAKING

by Susan J. Beck, Head of Public Services, Paul Robeson Library, Rutgers, the State University of New Jersey

As part of her sabbatical, Susan Beck served as an ARL Visiting Program Officer in 2002 to investigate the impact of assessment on library management decision making and the degree to which assessment data has influenced change. She presented papers on the results of her research at the ACRL 11th National Conference and the 5th Northumbria International Conference on Performance Measurement in Libraries.[1]

In this age of increased accountability, academic research libraries must effectively present information on their accomplishments to university administrators. Often, these administrators are looking for hard data as evidence of organizational improvement. In order to learn more about how data can influence decisions, nine ARL public universities actively involved with assessment activities were examined for patterns in their use of data to make decisions (Arizona, Arizona State, Connecticut, Texas, Toronto, Virginia, Virginia Tech, Washington, York).

Fifty-nine senior library managers and directors were interviewed about issues such as accountability, governance, service evaluation, planning, decision processes, assessment tools, and new data measures. Interview responses were analyzed using Atlas.ti: The Knowledge Workbench, a qualitative data analysis tool.

As expected, library administrators use data for planning and setting organizational priorities. Data are necessary for developing policies, allocating staff, building collections, and planning service delivery. The kinds of data used include formal assessments, survey results, and anecdotal information. Informed decision making lends credibility to both the library and its leadership. At some institutions the libraries have taken the lead in providing data, while at others the parent institution requires that all funding requests be accompanied by appropriate data. One library has a formal Management Information System by which it can generate the data necessary for units to support their planning processes. Some library administrators also feel that it is important to encourage data-driven decisions at all levels in the organization.

Nearly every unit in the library can use data for decision making. Collection development data for both print and electronic resources include usage data, survey data, economic information, departmental needs, shelving and interlibrary loan statistics, server logs, and information from faculty consultations. Some librarians have developed specialized local tools such as databases or Web sites. Special and digital collections data include information about users, research topics, and materials use. Usability testing of digital resources and server log analyses can help determine how those collections can best be used.

Human resources can conduct climate surveys to determine work satisfaction and identify problem areas. They often use data from focus groups, user (staff) surveys, and exit interviews. Analyses of peak service use can highlight staffing patterns to provide information for better scheduling. Process improvement studies can be used to discover areas for change in allocating staff resources.

Data are critical for making decisions about library space and equipment. Increased demand for instructional space points to the need for additional classrooms and sophisticated equipment. Performance of equipment can be tracked through logs and repair records. Student surveys often provide the needed data to support proposals for increased technology.

Decisions about services are influenced by user surveys, use data, focus groups, budget cuts, and process improvement activities. As an example, library service hours in all libraries visited changed as a result of data on use of services. Other services in which changes were made included interlibrary loan, desktop delivery and virtual reference services, shelving, and electronic reserves.

Results of the interviews indicate that assessment data provide necessary evidence to implement change in research libraries. Each visited library could point to several examples of the successful use of assessment data. Libraries that create, nurture, and integrate assessment into their everyday activities are models for other institutions. The willingness of leadership to incorporate assessment data into decision-making activities can be reflected in their own interest and the organizational culture, as demonstrated in staff expertise and concern for service quality. Accountability pressures from the parent institution are also a significant factor for these institutions and the use of assessment data provides the library with an opportunity to highlight its accomplishments with credibility and integrity.

—Copyright 2003 Susan J. Beck

[1] Susan J. Beck, "Making Informed Decisions: The Implications of Assessment," paper presented at ACRL 11th National Conference, Charlotte, N.C., April 10–13, 2003; and "The Extent and Variety of Data-Informed Decisions at Academic Libraries: An Analysis of Nine Libraries of the Association of Research Libraries in the U.S.A. & Canada," paper presented at 5th Northumbria International Conference on Performance Measurement in Libraries, Collingwood College, Durham, England, July 28–31, 2003.

ARL Service Quality Evaluation Academy

by Consuella Askew, LibQUAL+™ Program Specialist

On May 16, 2003, ARL's Service Quality Evaluation Academy graduated its second class of participants at the close of an intensive five-day workshop in San Antonio, Texas. The Academy emphasizes basic concepts and skills in measurement and data analysis for library service quality evaluation. The curriculum focuses on qualitative and quantitative methods with some time spent on relevant data-analysis software, such as ATLAS.ti and SPSS. The workshop is directed towards conducting service quality evaluations in general, with the expectation that participants will experiment with methods and eventually create their own tools to enhance the ones being developed by libraries.

The Curriculum

The first half of the workshop was devoted to quantitative methods and was led by Bruce Thompson, Professor and Distinguished Research Scholar from the Department of Educational Psychology at Texas A&M University. Participants benefited from class lectures as well as hands-on experience using SPSS software. Thompson introduced statistical terminology, rules for statistical analysis, and statistical methodologies to determine validity, reliability, and significance.

The second half of the workshop, led by Colleen Cook, Interim Dean of University Libraries, Texas A&M University, concentrated on qualitative research methodology. This segment of the program began with a half-day lecture on the use of qualitative methodologies. Then participants were introduced to focus group interview techniques. Participants were assigned to small groups and provided with interview transcripts, which they coded using ATLAS.ti and then discussed with the entire class.

Participant Feedback

Participants' reaction to the Academy was overwhelmingly positive, with many wishing they could devote more than one week to the event.

The ultimate goal of the Academy is to develop a critical mass of individuals with assessment skills within the library profession. A majority of the participants recommended that a follow-up training session would facilitate their use of the tools and methodologies taught in the Academy. They also indicated that they would enroll in an online course that enhanced what they learned in this workshop. As we listen to our participants, it is evident that, while the Academy is progressing toward its goal, there are additional learning opportunities to explore.

For information about participating in the next ARL Service Quality Evaluation Academy, contact Consuella Askew <consuella@arl.org>.

Creating a Culture of Assessment—An ARL/OLMS Workshop

Customer- or user-centered decision making and programming requires relevant data and information but also an organizational bias toward continuous assessment. The ARL Office of Leadership and Management Services (OLMS) offers a workshop on Creating a Culture of Assessment by investigating the main forces influencing the library and how these forces beg for the adoption of assessment and evaluation as an integral part of library work.

Workshop participants practice: integrating assessment work into organizational systems and structures and everyday work activities; determining current and preferred future paradigms and skills to move groups toward the new paradigm; and customer-focused data gathering.

For more information, including workshop schedule and registration form, see <http://www.arl.org/training/institutes/culture.html>.

Measuring Library Service Quality—An ARL/OLMS Online Lyceum Course

Measuring Library Service Quality is a Web-based course, offered over a six-week period by the ARL/OLMS Online Lyceum. The course covers the background and theory of measuring service quality, methods to assess and improve service, and the impact of measuring service quality on overall library value to constituencies. This course is designed for those interested in improving library service quality through a better understanding of how services are utilized, perceived, and assessed by users. This Online Lyceum course incorporates elements of both synchronous (real-time) and asynchronous interaction with course facilitators and a global peer network of up to 30 learners via a course bulletin board, chat rooms, and regular e-mail.

For detailed course information, including a course schedule, visit <http://www.arl.org/training/quality.html>.

Staying Current with the ARL Statistics & Measurement Program

To keep up to date with ARL Statistics and Measurement program activities, visit <http://www.arl.org/stats/>. You may also subscribe to the ARL new measures electronic mailing list. Members of the list are welcome to share ideas and information about performance measures, quality initiatives, and characterization of research libraries for discussion. See <http://www.arl.org/stats/newmeas/newmeas_list.html>.