Program Assessment in Academic Libraries: An Introduction for Assessment Practitioners

Eric Ackermann
Radford University

Abstract

Although academic libraries have a long tradition of program assessment, in the past the results have been more meaningful internally than externally. Recent changes in the conceptualization of libraries' role in higher education and advances in measurement tools will likely provide answers to different questions, particularly the relationship of library services and resources to student learning and success.

Introduction

The academic library sits at the intersection of university instruction, services, and resources. It is not an academic department, yet it provides instruction on information literacy. It is not part of student support services, yet it provides student services such as research help, coffee shops, and study spaces. It is not academic computing, the bookstore, or facilities management, yet it provides similar resources (e.g., computers, books and journals, and a comfortable place for study and reflection). Traditionally, librarians assumed that all these components contributed to student learning and student success. Consequently, for years academic libraries have documented and used assessment data focused on the quantity and utilization of resources.

However, in the context of contemporary institutional assessment, this type of data is not sufficient. External stakeholders now question the link between resources and learning outcomes, no longer taking the previously assumed relationship for granted. While information about resources and resource utilization is undeniably important, it fails to address exceptionally important questions. For example, how specifically are academic libraries contributing to student learning outcomes? How are these contributions measured in ways that are meaningful to stakeholders outside of the library world? To explore these issues, this paper provides an introduction to academic library programs and what they assess.

For the sake of clarity, I will define a few terms that occur repeatedly in the text. *Academic libraries* are "libraries in higher education, from technical institutes to research universities" (ACRL, 2004). *Research libraries* are a sub-group of academic libraries that are associated with comprehensive, doctoral granting, research universities (ARL, 2006a). Traditional library metrics focus on *inputs* or "the raw materials of a library program-the money, space, collection, equipment, and staff," and *outputs* or measures that "quantify the work done, i.e., number of books circulated, number or reference questions answered." Current library assessment focuses more on *outcomes* or "the ways in which library users are changed as a result of their contact with the library's resources and programs" (ACRL, 2004). Some of the current library assessment tools are only *data collection tools*. Other tools are *data collection and*

analysis tools that provide both data collection capabilities and a descriptive statistical analysis of the results, which often includes raw data and some level of comparative data.

The balance of this paper is organized into three sections (1) the tradition of assessment in libraries (2) the current state of affairs and challenges of assessing the following library components: instruction, services, and resources and (3) implications for the future of library assessment.

Traditional Library Assessment

Academic libraries in the United States have a long tradition of assessment. It began in 1906 when James Gerould at Princeton started collecting library statistics for selected college and university libraries. Later these libraries formed the Association of Research Libraries (ARL) in 1932. In 1961 ARL took over the gathering and annual distribution of statistics on behalf of its 123 member libraries in the United States and Canada (ARL, 2006a; ARL, 2006b).

The distinctions between research libraries and non-research libraries sharpened during the early twentieth-century. By 1979 the Association of College and Research Libraries (ACRL) (established in 1890) began systematically collecting and publishing statistics for the academic libraries not covered by the ARL in *University Library Statistics*. In 1998 this effort evolved into *Academic Library Trends and Statistics*, which provides annual data from 1215 academic libraries in the US and Canada (ACRL, 2003).

Both the ARL and ACRL statistics focus on gathering similar types of input and output data for comparative purposes over time (see Table 1). It has been assumed that if the academic libraries provided these inputs and outputs, the desired outcomes such as student learning and research success will automatically follow (Kyrillidou, 2002).

Table 1

Traditional Library Data Collection Tools

Name	Measurement focus	Data categories and format	Results reporting	Availability (contact information)
ACRL Academic Library Trends and Statistics	Input and output variables of 1215 academic libraries from all Carnegie classifications; in US and Canada; coverage from 1998 to present	6 core data categories (Collections; Expenditures; Electronic Expenditures; Personnel & Public Services; PhDs granted, Faculty & Enrollment; Networked Electronic Resources & Library Digitization Projects); web delivered or print	Category by Carnegie classification	Annually, by subscription (acrl.telusys.net/trendstat/2005)
ARL Statistics	Input and output variables of 123 member academic libraries in US and Canada; coverage from 1906 to present	7 data categories (Library characteristics, Collections, Personnel and Public Services, Expenditures, Electronic Resource Expenditures, Service Items, University Data); web delivered or print	Reported by member institution; ranked results by category	Annually, by subscription or purchase (www.arl.org/stats/arlstat/index.html)

In the last two decades, the implicit relationship among inputs, outputs, and student learning has been under fire by external stakeholders such as regional accreditors and other governmental bodies. Reflective of this heightened scrutiny, library assessment is trending

toward student learning outcomes and how inputs and outputs relate to them. The intent is to capture the academic library's contribution to institutional effectiveness and student learning outcomes (ACRL, 2004). The next section focuses on the current state of assessment in three important components of libraries: instruction, services, and resources.

Current Library Assessment

Instruction

The goal of a library instruction program is the teaching of information literacy (IL). IL is related to but distinct from information technology (IT). IT skills represent facility with technology itself. IL, on the other hand, is an "intellectual framework for understanding, finding, evaluating, and using information." Students may use IT in the process of demonstrating IL, but proficiency is mainly dependent on the use of "critical discernment and reasoning" (ACRL, 2000). Types of IL instruction include hands-on active learning workshops on using online databases, avoiding plagiarism, and evaluating websites. IL instruction may also include course related, group instruction in e-classrooms and web tutorials.

Almost from its inception attempts have been made to assess the efficacy of library IL programs (Bober, Poulin, & Vileno, 1995; Rader, 2002). Until recently, however, virtually all the instruments geared toward assessing IL programs were locally developed. Unfortunately, these early instruments yielded results of dubious reliability and validity. Efforts to address these shortcomings led to the development of several new assessment tools, such as Project Standardized Assessment of Information Literacy (SAILS), iSkills, James Madison University's (JMU) Information Literacy Test, and the South Dakota Regental Information Literacy Exam (SDILES) (see Table 2). Though different, they are all data gathering and analysis tools, normed at the state or national level, and are based on one or more parts of ACRL Information Literacy Competency Standards for Higher Education (ACRL, 2000).

Table 2

Data Collection and Analysis Tools for Assessing Library Instruction and Information Literacy Programs

Name	Measurement focus	Format	Data analysis	Results reporting	Availability (Contact Information)
Project SAILS (Standardized Assessment of Information Literacy)	IL skill sets based on ACRL standards 1, 2, 3, & 5 for student cohorts (groups)	45 multiple choice questions randomly chosen from test bank of 250; Web delivered, paper optional	Item response theory (reported reliability = .80); average student cohort performance	By ACRL standard & skill set; by demographics, class standing	Open: Fall 2006 & Spring 2007 (www.projectsails.org)
Information and Communication Technology (ICT) Literacy Assessment	Critical-thinking via 7 ICT skills based on ACRL standards 1, 2, 3, & 4	Real-time performance on 14 four minute & one 15 minute tasks; core & advance versions; web delivered	Individual scores; performance feedback by task	Overall scores compared to other test takers; performance feedback by task	Open: anytime (www.ets.org/ ictliteracy)
James Madison University (JMU) Information Literacy Test	Knowledge & application of IL skills based on ACRL standards 1,2,3, and 5	60 multiple choice items: 41 knowledge & 19 application	Reported reliability = .88	Provides data set of scores & scored responses	Open: anytime? (www.jmu.edu/icba/ prodserv/ instruments_ilt.htm)
South Dakota Regental Information Literacy Exam (SDILES)	ACRL based IL skills at document (minimum) level & assessment (continuous) level	25 multiple choice questions randomly chosen from test bank @ ratio of 3 documentation to 2 assessment items; web delivered via WebCT	Item response theory; 2 scores per individual: documentation (pass/fail, cut score 13/25) & assessment (per ACRL standard)	Assessment scores for library; document scores for university admin and state legislature	In development, recruiting participating institutions (Carol Leibiger, C.Leibiger@usd.edu; William Schweinle, (William.Schweinle@ usd.edu)

Measuring the effectiveness of IL instruction can be challenging. While some IL instruction is course integrated or is implemented in a one-hour credit general education course, most are conducted in single session classes at the request of a faculty instructor (Wong, Chan, & Chu, 2006). There are no stable groups of students to access, and the scope and content of each session can vary widely, depending on the needs of the faculty requesting the session. Because of the variability in how IL instruction is implemented – particularly via short sessions - most libraries have had difficulty demonstrating its affect on student learning and success.

Services

Academic libraries provide a range of services, including convenient access hours, public access computers, wireless Internet access, coffee shops, individual and group study spaces, interlibrary loan and document delivery, and the circulation of books, dvds, and other materials. In addition reference services provide research and technical support that involves personalized instruction for identifying and locating research materials and answering a myriad of specific questions (ACRL, 2004).

Service quality assessment is based on the assumption that customer perceptions are a valid source of information about the type and quality of services provided (Kyrillidou & Heath, 2001). One of the earliest and most widely used instruments is LibQUAL+ (see Table 3). It evolved from the SERVQUAL+ customer service instrument used in the private sector (Parasuraman, 2002). Libraries use information obtained by this instrument to modify services in response to changing customer needs (Heath, Kyrillidou, & Askew, 2004). Nowhere do customer needs change more rapidly than in the online, digital environment. To address this virtual area, an online service quality assessment tool, DigiQUAL, is being developed by the makers of LibQUAL+ (ARL, 2005; see Table 3).

Table 3

Data Collection and Analysis Tools for Assessing Library Service Programs: Library Service Quality

Name	Measurement focus	Format	Data analysis	Results reporting	Availability (contact information)
LibQUAL+	User's perception of library service quality across 3 dimensions (Affect of Service, Information Control, Library as Place)	Survey; 22 items & 1 comment box; each rated on 9-pt Likert scale for min. & max. expectations, & current satisfaction; web delivered	Gap analysis; average scores by group (faculty, staff, students; reported reliability > .80; comment data unanalyzed	Average scores by item and dimension for each group; comment data by individual respondent as text file	Open: Fall & Spring each year (www.libqual.org)
DigiQUAL	User's perceptions of digital library website quality (reliability, functionality, content)	Survey; 5 items chosen from bank of 180 items, 1 fixed item, & 1 comment box; each item rated on 7-pt Likert scale for personal importance & site performance; web delivered	Adaptation of LibQUAL+ protocol	Report of results by item	In development, recruiting participating institutions (www.digiqual.org/ digiqual/index.cfm)

Traditional reference statistics track the number of reference transactions by date/time and mode of delivery such as walk-up or phone, and are often accompanied by daily logs of unusual, ongoing, or complex transactions. These statistics are used mainly to inform staff management decisions and for reporting to ACRL and ARL. This pencil and paper system is

clumsy, time-consuming, inconsistent, unpopular, and too narrowly focused to capture the new ways reference transactions are happening (McClements, Vack, & Calcese, 2005; Smith, 2006). Fortunately, several promising tools are either available or in development to address these shortcomings such as LibStats, the WCL Reference Statistics System, and Desk Tracker (see Table 4). All of these tools capture the new modes of reference transactions such as email and instant messaging, and previously unrecorded transactions occurring at diverse, non-traditional locations (e.g., in one's office or at the circulation desk). The developers of LibStats and the WCL Reference Statistics System are also considering further uses of data to enhance marketing, resource development, and instruction (McClements, Vack, & Calcese, 2005; Smith, 2006).

Table 4

Data Collection Tools for Assessing Library Service Programs: Reference Services

Name	Measurement focus	Format	Data analysis	Results reporting	Availability (Contact Information)
LibStats	Reference services: user counts & reference transactions over time; not per ARL standards	3 types of data categories: transaction type, date/time of transaction, total transactions; stored in web accessed database	None provided	Report generator as Excel spreadsheets; report type: questions by date, patron type, format, time of day, weekday; or all data	Available as freeware (www.wendt.wisc.edu/projects/systems/libstats/) or contact Nathan Vack, njvack@wisc.edu)
WCL RefStats	Reference services: hourly user counts, & reference transactions over time; per ARL standards	4 types of reference transactions (directional, technical, ready ref, other/consultation); accessed via web; data stored as Access database record	None provided	Reports available by service provider/ librarian, transaction type (or source), user type (or department); daily activity log; custom/ unique reports	In development. Source code available. (Michael Smith, michaelsmith @tamu.edu)

Currently service assessment is based primarily on customer (student, faculty, and staff) satisfaction survey instruments. Such measures, however, can be of limited utility to academic libraries for funding. In the highly competitive world of research universities, all units are fighting for students, faculty and money (Lombardi, 2006). Even if data suggest that students, parents, and alumni love the library, they will probably like other areas better. For example, Lombardi (2006) pointed out that if projects were approved based solely on client satisfaction, then a new or renovated library would stand little chance of funding compared to a new sports/entertainment complex. For funding decisions libraries should pitch their needs in terms of how they will affect student outcomes and success, and how these needs are central to their respective universities' missions.

Resources

Within budget constraints, library resource programs provide access to a wide range of authoritative and up-to-date resources in diverse formats that support the curriculum and the needs of its users (ACRL, 2004). Libraries provide access to these resources either directly through physical ownership of books or indirectly via subscription to online e-resources (e.g., databases such as InfoTrac).

Traditional resource metrics include usage statistics, such as tracking how often books are checked out and in-house reshelving counts for journals and reference books. As an increasing number of resources are accessible only online, traditional print based statistics are increasingly unrepresentative of available library resources.

Vendor supplied and web log use statistics for e-resources help supplement the traditional metrics. However, the utility of vendor statistics is hampered by inconsistent measurement frames, differing metrics, and different definitions for the same metrics (COUNTER, 2007). Web logs suffer from a lack of granularity, standardized metrics and reporting protocols that allow comparison among institutions. More importantly, neither method captures the "why," or the purpose of the use. Without this information, it is virtually impossible to determine if the resources are being used to advance student learning or any other desirable outcome. One solution in development by ARL is the MINES for Libraries protocol. It is an online "transaction-based survey that collects data on the purpose of use of electronic resources and the demographics of users" (ARL, 2005; see Table 5).

Table 5

Data Collection and Analysis Tools for Evaluating Library Resource Programs

Name	Measurement focus	Format	Data analysis	Results reporting	Availability (Contact Information)
MINES for Libraries (Measuring the Impact of Networked Electronic Services)	Electronic resource usage (purpose)	Transaction based usage survey; web delivered each time an e-resource accessed; 5 items & 1 comment box	Not known	Summary tables of frequency (use) counts by user group, discipline, location, use purpose, reason for e-resource selection	In development (www.arl.org/stats/newmeas/mines.html)

The challenge facing resource assessment is two-fold. It still must demonstrate return on investment and accountability despite budgets that fail to keep pace with rising costs. Resource assessment must also incorporate methods for determining and demonstrating the impact or links between resource use and desired outcomes, such as student learning and success.

Implications for the Future of Library Assessment

Academic libraries have a tradition and culture of assessment stretching back almost one hundred years. This history is both an advantage and a disadvantage. On the positive side, the culture already exists; librarians are accustomed to tracking certain types of information. The primary disadvantage is that most library assessment is developed in relative isolation from the larger higher education community. It has been driven mainly by internal library needs, and has resulted in metrics and reporting protocols that are meaningful primarily to other librarians.

Instead, these measures need to be meaningful not only to librarians but also to the other stakeholders, both on campus (e.g., institutional research, university administrators), and off-campus (Lombardi, 2006; Leibiger & Schweinle, 2006). Fortunately, developing such measures is not an insurmountable problem as demonstrated by the success of the SDILES at simultaneously producing IL assessment results meaningful to all three of the main stakeholders: the library, the university, and the state (Leibiger & Schweinle, 2006). Much work has yet to be done in the assessment of libraries, but hopefully these future endeavors will illuminate relationships between elements of the library and student learning and success. Such information could only help the efficacy of libraries around the country.

References

Association of College and Research Libraries. (2004). Standards for libraries in higher education. Retrieved January 1, 2007, from http://www.ala.org/ala/acrl/acrlstandards/standardslibraries.htm

Association of College and Research Libraries. (2003). Chapter 15: History. Retrieved December 31, 2006, from http://www.ala.org/ala/acrl/aboutacrl/resourcesforwork/guidetopolicies/chapter15history.htm

Association of College and Research Libraries. (2000). Information literacy competency standards for higher education. Retrieved January 1, 2007, from http://www.ala.org/ala/acrl/acrlstandards/informationliteracycompetency.htm

Association of Research Libraries. (2006a). About ARL. Retrieved January 1, 2007, from http://www.arl.org:16080/arl/

Association of Research Libraries. (2006b). ARL statistics. Retrieved December 31, 2006, from http://www.arl.org/stats/arlstat/

Association of Research Libraries. (2005). MINES for Libraries, Measuring the Impact of Networked Electronic Services. Retrieved January 14, 2007, from http://www.minesforlibraries.org/

Bober, C., Poulin, S., & Vileno, L. (1995). Evaluating library instruction in academic libraries: A critical review of the literature, 1980-1993. *Reference Librarian, 51/52,* 53-71.

COUNTER, Counting Online Usage of NeTworked Electronic Resources. (2007). About COUNTER. Retrieved January 14, 2007, from http://www.projectcounter.org/about.html

Heath, F., Kyrillidou, M., & Akew, C. (Eds). (2004). *Libraries Act on Their LibQUAL+ Findings, From Data to Action.* Binghamton, NY; Haworth Press, Inc.

Kyrillidou, M. (2002). From input and output measures to quality and outcome measures, or, from the user in the life of the library to the library in the life of the user. *Journal of Academic Librarianship, 28,* 42-46.

Kyrillidou, M., & Heath, F. (2001). Introduction: Measuring service quality. *Library Trends, 49,* 541-547.

Leibiger, C., & Schweinle, W. (2006, April). *The South Dakota Regental Information Literacy Exam: A tool for small and medium-sized universities to document and assess information literacy.* Presentation presented at the North Carolina State Undergraduate Assessment Symposium, Cary, North Carolina. Retrieved January 1, 2007, from http://www.ncsu.edu/assessment/symposium/presentations/A36.pdf

Lombardi, J. (2006, September). Library performance measures that matter. Presentation presented at the Library Assessment Conference, Charlottesville, Virginia.

McClements, N., Vack, N., & Calcese, S. (2005, April). Converging statistics. Presentation presented at the Wisconsin Association of Academic Librarians Conference, Madison Wisconsin.

Parasuraman, A. (2002). Forword. *Performance Measurement and Metrics, 3*, 37-39.

Rader, H.B. (2002). Information literacy, 1973-2002: A selected literature review. *Library Trends, 51*, 242-259.

Smith, M. (2006). A tool for all places: A web-based reference statistics system. *Reference Services Review, 34*, 298-315.

Wong, G., Chan, D., & Chu, S. (2006). Assessing the enduring impact of library instruction programs. *Journal of Academic Librarianship, 32*, 384-395.