

*“The Louder they Talked of Outcomes, the Faster We Counted Our
Beans”*

MEASURING THE IMPACT OF ACADEMIC LIBRARIES

SCONUL, LONDON 29TH NOVEMBER 2009

*MICHAEL GORMAN
DEAN OF LIBRARY SERVICES
CALIFORNIA STATE UNIVERSITY, FRESNO*

Institutions of higher education have three fundamental purposes—to educate their students, to facilitate the research activities of their faculty members, and to serve the community and the region in which they are situated. These widely accepted activities are open to much interpretation and many questions. What are the characteristics of an educated graduate? What is the value of pure research and of applied research? How do we measure those values? What is the community service role of the university and of its students and faculty? Academic libraries serve their universities and colleges by supporting teaching/learning, research, and service by giving the academic community access to, and assistance in the use of, the human record—both that sub-set of the human record owned by the library and the wider human record in the libraries of the world. Given that role, it is obvious that we should judge the value and outcomes of academic library service in the light of the general purposes of the academic institutions they serve.

North American academic libraries, indeed all of academia, are awash with talk of “outcomes,” “assessment,” and “accountability.” They are the buzziest of buzzwords and have quite replaced all the terms derived from Japanese business management that held sway as recently as 15 years ago. There is a discontinuity at the heart of the outcome of all the talk about outcomes that inspires me to paraphrase Ralph Waldo Emerson—“The louder they talked of outcomes, the faster we counted our beans.”ⁱ Higher education is complex, various, and multi-faceted—it can only be understood in the light of sophisticated, nuanced evidence gathering. The products of bean-counting do not enable us to understand, nor does the application of the methods favoured by Mr. Gradgrind, that “eminently practical man” who believed in a life governed by facts and statistics and in the ruthless suppression of imagination and the arts.ⁱⁱ

The evaluation paradox

A leading US academic librarian has written:

The future vitality of libraries in academia will be dependent on whether they can ... prove their value to the overall academic endeavor. This value must be documented at a level that transcends specific formats of information, [and] locations of collections and location of users, and that clearly links the investment in campus-wide information resources to the effectiveness of particular disciplinary programs.ⁱⁱⁱ

If we are to assess the impact of our libraries, we must engage in evaluation leading to an understanding of the outcomes of what we do. One of the very first classes that I taught at the University of Illinois library school thirty years ago was “The measurement and evaluation of library services” based on the manuscript of the first edition of the book of the same name by fellow-Briton Wilf Lancaster.^{iv} This brilliant synthesis of the studies of, and research into, reference services, collections, catalogues, etc. of libraries both impressed me with its scope and presented me with a paradox that continues to confound me after all these years. That paradox can be summed up thus: it appears that the only evaluative results that can be accepted with confidence are those that pertain to the most banal aspects of library services. Conversely, the more valuable evaluative data is, the harder it is to arrive at them. It is comparatively easy to gauge the accuracy of an informational reference question (“What is the date of the New Hampshire primary election in 2008?”) but much more difficult to evaluate the worth of a half-hour consultation with a reference librarian on a thesis topic. “Does your catalogue bring together all the works of one author?” can readily be tested and answered. “What is the value to the library’s users of the elaborate syndetic structure of the catalogue?” is a much thornier question. This paradox applies, of course, to all aspects of life and to all institutions. You can guess that someone is a good Catholic by observing that she goes to Mass every Sunday and takes the sacraments at

appropriate intervals. The more important questions of faith, intentions, and works are often not susceptible to such observation and enumeration.

Tests versus evaluation

It may be simplistic to say that, if you can count something easily, the results will not be very informative, but it is true that effects of importance are almost invariably very difficult to measure. As I have said, in education (in both schools and academia) in the US and, for all I know, elsewhere, the talk is all of “outcomes” and “accountability.” These words pervade all exercises such as accreditation, self-studies, and, ultimately, budgeting. In the schools (kindergarten through 12th grade in American terms) the “No Child Left Behind” Act holds sway.^v This unprecedented incursion by the federal government into what has hitherto been the domain, and funding responsibility, of state and local governments depends almost entirely on standardized testing of reading, writing, and mathematics at the various grade levels. California, the state in which I live, has introduced a standard test that all students have to pass before they can receive a high-school diploma, irrespective of their grades (marks) and achievements during their four years in high school.^{vi} Though both of these measures have been heavily criticized—the federal act on a number of grounds including the usual “unfunded mandate” question—few have questioned the idea of ensuring that the considerable sums spent on education are accounted for. The trouble lies not

in the ideas of assessment and accountability but in the methods used—in this case standardized tests. The reason why many object to these tests is that they are one-dimensional pictures of achievement that do not take more nuanced assessments into account and that only reward minor attributes—the ability to take tests, memorize facts, and guess the correct line in a multiple choice test. The unfortunate thing is that those tests are there because they are the easiest way to measure, or appear to measure, ability. They are chosen in both the NCLB regulations and the California high school graduation requirement because they are easy to administer and to mark not because they address real achievement or educational progress. Those things are beyond the powers of the federal government or the state—this forces the bureaucrats and politicians to pretend that the things that they can measure are as important as the things they cannot. It makes us all complicit in, essentially, a fraud and pretence.

Hard questions

The fruits of education are complex and multi-dimensional as are the fruits of good library service. It behooves us avoid the mistake that has been made elsewhere in education and to ask hard questions. How do you measure the value of use of the library and its services? How do you arrive at the outcomes of access to a major library collection built up over many years and a wide range of digital resources? How do you assess the value of the library's services to faculty research? What is the measurable value of a library that

defends the intellectual freedom and privacy of its users? Have university libraries an economic value to their institutions? Can non-profit cultural institutions such as libraries use assessment techniques that are developed for commercial for-profit enterprises? What is the value of the academic library as centre of the intellectual life of the university?

Characteristics of papers on evaluation

I have read a large number of articles and papers written on these topics as they apply to North American university libraries in the last decade or so. These documents share a number of characteristics (apart from being written in an almost uniformly execrable and jargon-ridden style).

- To a great extent, they concentrate on digital resources and services.

This is partly due to the, alas, endemic technophilic faddishness and to the political pressures exerted on the library by its masters but also to the fact that quantitative measures of use of digital resources are far easier to gather than those relating to other library services.

- They share a concern about the imprecision of terminology and concepts. One important paper asks:

What is the difference between performance measures and indicators? What is meant by “valued institutional outcomes? Is the evaluation of library effectiveness the same as library quality or performance?^{vii}

This concern about specific terms is but one example of the general confusion about the meaning of an “outcome;” the definition of certain kinds of outcome; the overlap between inputs, outputs, user satisfaction, and outcomes; and many other questions of definition. McClure and Lopata give one useful set of definitions:

Performance measures are a broad managerial tool that encompass [*sic*] measures of *inputs* (indicators of the resources essential to provide service); *outputs* (indicators of the services resulting from the use of those resources) and *impacts* (the effect of those outputs on other variables or factors) [*emphases added*].^{viii}

- There is a lack of standardization between libraries when it comes to their assessment procedures and measures and the “lack of performance measures that make sense across institutions and that link library processes to educational and research outcomes.”^{ix}
- The idea of aligning the library’s outcomes with the mission, goals, and performance measures of the parent institution presents inherent difficulties at many levels—terminology can differ widely; the goals and mission of the university as expressed in academic strategic plans are often geared to the schools, colleges, and departments and may be difficult to match with, or only obliquely related to, those of the library; and, the goals and mission of the of the university are often expressed in cloudy language signifying imprecise thought and conflicting visions.

There are, however, some areas that many writers feel are pertinent to the alignment between the goals of the university and the goals of the library. In the view of one writer, these include:

... measures of academic library use and library skills instruction correlated to lower attrition rates, higher grades, higher GRE scores, student persistence, and savings in faculty time.^x

Understanding the limitations of evaluation is a necessary precondition of successful evaluative processes. Rather than bewailing those limits, we should see them as liberating in that they lead to both realistic expectations and productive assessment procedures.

A culture of evaluation

All too often, the honest answer to the question “why are you doing evaluation of your library?” is “because we are required to do evaluation by our masters.” This is the worst of all reasons and I would urge us all to get past the fact that evaluation is mandated and seek other better reasons. Most North American universities are striving to establish what they call “a culture of evaluation.” This culture has, as one of its fundamental characteristics, the idea that such evaluation is continuous and cumulative—in other words, evaluation should be and integral and permanent part of every programme and service. Other characteristics of the culture of evaluation are, in the words of a paper on the topic:^{xi}

When people claim that their organisation has an evaluation culture

... they usually mean that all members of the organisation:

- accept the use of evaluation;
- understand why the organisation uses evaluation;
- can design or get advice on design of necessary evaluations;
- use evaluation, particularly to support change and development.

In other words, they refer to a known, shared policy about evaluation within the organisation.

This transformation of a way of looking at our work is, possibly, more difficult than carrying out evaluations themselves, but, if achieved, makes those evaluations and the search for outcomes easier.

In addition to accepting the psychological reorientation needed to create a culture of evaluation, success depends on understanding the characteristics of successful evaluation. These characteristics are, in the words of two American authors:^{xii}

1. “Evaluation results from design, not accident.” By this they mean that evaluation must proceed from planning and the planning should have specific aims.
2. “Evaluation has a purpose.” By this they mean that successful evaluation depends on clear goals and objectives, clearly enunciated and communicated to all involved in the process.

3. “Evaluation is about quality.” By this they mean that library evaluation must have an over-riding aim—that of improving the quality of library services and programmes. Another author has written, “Service quality has traditionally been viewed in terms of gap analysis—the extent of the gap between ideal and actual service expectations.”^{xiii}
4. “Evaluation is more than measurement.” By this they mean that measurement is a tool in evaluation but that measurement in itself accomplishes very little unless it is a part of a well thought out plan with useful objectives..
5. “Evaluation can be small-scale.” By this they mean that a single, brief evaluative exercise, if properly planned and executed and if dedicated to a useful purpose, can be both valuable and cost-effective.
6. “There is no one right way to evaluate.” By this they recognize that libraries differ in mission and purpose, serve different institutions, and possess their own values and cultures. That being so, there is no best practice in evaluation, just a commonality of approach based on an understanding of the characteristics of successful evaluation.

The authors then go on to give reasons why libraries should create a culture of evaluation. These are, essentially, to solve problems; to help in administrative decision-making; for public relations purposes; and, to be successful politically.^{xiv} All libraries have problems—collections and services

that could be better, personnel and workflow problems, use of outdated physical plant, etc.—and those problems are only solvable if understood. The evaluative process enables definition and understanding and leads to possible solutions. Evaluation is an essential part of efficient administration because it leads to accurate information on which to base change and improvement. Public relations and politics are entwined inextricably and the modern library has to be good at both.

Assessing the value of academic libraries

There have been a number of well-regarded studies demonstrating the economic value of public libraries to their communities.^{xv xvi} All those studies agree that a dollar invested in a public library will produce many dollars in return. Other studies have shown the direct benefit in terms of student achievement of well-funded and staffed school libraries and media centers, notably those undertaken by Keith Curry Lance and his colleagues at the Library Research Center in Colorado.^{xvii} It is, perhaps, understandable that few such studies exist when it comes to college and university libraries. Public libraries at the centre of thriving communities can readily be shown to have a direct economic benefit to those communities, as well as the less tangible benefits to society. School libraries and media centres exist in the relatively simple and structured environment of elementary, middle, and high schools. The measurement and assessment of the economic and scholastic benefit of

both libraries is a comparatively simple task. On the other hand, there are enormous disparities between the libraries of community colleges, private liberal arts colleges, state universities great and small, and the major public and private research institutions of North America—in addition, the multi-dimensional environment of each of these institutions makes meaningful assessment of outcomes very difficult. In the words of one accrediting agency:

... every institution must be considered within its own setting and not by comparison with general patterns or norms.^{xviii}

User satisfaction surveys

The Association of Research libraries has developed a set of assessment tools under the name of LIBQUAL+™, which is used to solicit, track, understand, and act upon users' opinions of library service quality. The programme is based on a Web-based survey linked with training.

LIBQUAL+™ is aimed at helping libraries to assess and improve library services, change organizational culture, and market the library. Its stated goals are to:

- Foster a culture of excellence in providing library service
- Help libraries to understand user perceptions
- Collect and interpret library user feedback systematically

- Provide libraries with comparable assessment information from peer institutions
- Identify best practices in library service
- Enhance the analytical skills of library staff ^{xix}

These worthy (and overlapping) aims are, I believe, overstated, because LIBQUAL+™ and other measures of user satisfaction are, ultimately, measuring subjective reactions and perceptions rather than objective outcomes—versions of reality rather than reality itself. I suspect they are popular because such surveys almost invariably grant high scores to libraries and their services (especially when libraries are compared to other service units on campus). Though they are valuable, for that reason, to the public relations and marketing aspects of the use of assessment, they are unlikely to suggest positive real world changes.

Areas of academic library assessment

Given the characteristics of successful assessment given above^{xx} and, in particular, the emphasis on planned, purposeful evaluation of the quality of library programmes and services, I would suggest that the following are important areas of the academic library that could be subject to a culture of evaluation.

- Quality of collections and electronic resources made available by the library. Such evaluations are either collection centered or user centered. Strategies that can be employed include mapping collections and electronic resources to curricula and to the agenda of research centres; the “could this book have been written here?” test; and other more sophisticated collection analysis procedures.^{xxi}
- Reference services. One interesting multi-tool approach to the assessment of reference services was taken at the University of Arizona libraries. Their assessment was based on a redefinition of the traditional role of reference librarians.

Librarians at the University of Arizona Libraries strive to help students gain user sufficiency and information literacy. When on the reference desk, they tend to teach rather than do the work for the students. As the libraries are customer centered, it was decided to conduct an evaluation of its reference services. The evaluation involved using a combination of surveys, focus groups, and unobtrusive observation worksheets with a small sample of students. The results were very meaningful. The unobtrusive observation of the reference interaction between student and reference provider gave the most useful results for improvement.^{xxii}

Another popular reference services assessment tool is called the Wisconsin-Ohio reference evaluation (WOREP). More than 100 academic libraries have used this tool, which was developed in the 1980s.

WOREP is designed to assess the outcome of the reference transaction and to identify factors related to success or lack of success. Its key features are:

- demonstrates validity--items actually measure what is indicated;
- demonstrates reliability--it is possible to replicate results;
- uses the natural reference center setting;
- includes sufficient factors to suggest cause and effect relationships;
- provides for comparability of data.

WOREP is unique in that it does not rely solely on patron input, but also collects data from staff. Also, it does not limit evaluation to patron satisfaction, but also considers the success of the transaction.

- Catalogue use studies. Thousands of catalogue use studies of varying utility have been published in the last decades. Many of them have been driven by the changes in public catalogues wrought by MARC, cataloguing codes and standards, the globalization of cataloguing data, and, latterly, the move of catalogues to the Web. Significant contributions have been made in studies of subject access, such as those by Karen Drabenstott and others.^{xxiii} Many other studies are banal and/or of local interest. The studies of metadata (a limited form of cataloguing for electronic resources) have, thus far, been characterized by polemics and assertion rather than meaningful data.
- Access, availability, and use. The question of the level of access to the library and its services (on-site and remotely) is central to an understanding of its value. The adequacy and inadequacy of the facility and the question of “the library as place” are important variables as are the equity of provision of services to all students, faculty, and other

library users, including remote users such as those engaged in distant learning.

- Service quality. The most obvious change in evaluation of both universities and their libraries is the move away from quantitative measures (except, interestingly, in evaluating the use of electronic resources) and toward subjective evaluation of quality derived from business practices.^{xxiv} The most widely used instrument for measuring service quality is called SERVQUAL, which seeks to “assess customer perceptions of service quality in service and retail businesses.”^{xxv} The scale measures service quality in five categories:

- Tangibles - physical facilities, equipment, staff appearance, etc.
- Reliability - ability to perform service dependably and accurately
- Responsiveness - willingness to help and respond to customer need
- Assurance - ability of staff to inspire confidence and trust
- Empathy - the extent to which caring individualized service is given

SERVQUAL, which was developed for a wide range of service industries, measures service quality as the discrepancy between a customer's expectations of a service and the customer's perception of the service received and, therefore, requires respondents to answer questions about both their expectations and their perceptions. The important things to note about SERVQUAL are, first, that it was developed for business—that is, it deals with

the bottom line—ultimately the sole measure of commercial transactions; and, second, that it measures perceptions not actuality.

Library instruction outcomes

Apart from the use of electronic resources, the most popular area for outcomes assessment in US academic libraries is the complex of activities variously called “bibliographic instruction,” “information competence,” and “library instruction,” in which librarians seek to teach undergraduate students the skills needed to locate and use relevant library materials and resources and the critical thinking skills necessary to evaluate those materials and resources. Because these activities are measurable in the same manner as classroom instruction elsewhere in the university, they fit into campus-wide measurement and evaluation. Many academic libraries publish the “learning outcomes” and how the students’ progress is to be assessed.^{xxvi} A typical statement of these outcomes is (from the Oregon State University library):

An OSU graduate:

Recognizes when information is needed. Finds information efficiently. Learns from information gathered. Uses information effectively and ethically.^{xxvii}

The California State University system has created a set of “Ten Basic Information Competency [sic] Skills:”^{xxviii}

1. Define the research topic.

- a. Formulating a research question.
- b. Selecting, refining/narrowing, a topic.
- c. Selecting search terms (keywords).
2. Determine the information requirements for the research question.
 - a. Steps of the library research process.
 - b. Controlled vocabulary (e.g. descriptors, subject headings).
 - c. Selecting appropriate sources (print and/or electronic).
3. Locate and retrieve relevant information.
 - a. Ask for assistance at the Reference Desk.
 - b. Refine vocabulary terms used.
 - c. How to differentiate between popular magazines and scholarly journals and why this is important.
 - d. How to differentiate between an electronic database and a web site and why this is important.
 - e. How to use the library catalog (author/title/keyword/subject restrictions).
 - f. How to read a call number.
 - g. How to read a citation (including the components of a citation).
 - h. How to systematically gather information.
 - i. Borrowing privileges at this library.
 - j. ILL.
 - k. Reserves.
 - l. Library locations.
4. Use the technological tools for accessing information.
 - a. Accessing the library databases from home.
 - b. NIS accounts.
 - c. Emailing, printing, downloading.

5. Evaluate information.
 - a. Critical thinking.
 - b. Critical evaluation of information found on the WWW.
6. Organize and synthesize information.
[not generally related to library instruction]
7. Communicate using a variety of information technologies (creating web pages, PowerPoint presentations, spreadsheets). [not generally related to library instruction]
8. Understand the ethical, legal, socio-political issues surrounding information and information technology.
 - a. Plagiarism.
 - b. Copyright.
9. Use, evaluate, and treat critically information received from the mass media. [not generally related to library instruction]
10. Judge the product and the process. [not generally related to library instruction]

(Note that only the first five steps are generally considered to be part of the library's instruction programme.)

One problem with all studies of information competence is that they lack definitions of “information” and, hence, any satisfactory assessment of the outcomes of competence in locating, evaluating, and using that elusive substance. For example, a basic outcome of higher education is the ability to interact with complex texts productively—to read a literary or scholarly book

and thereby gain knowledge. How is that fundamental ability to be measured? Another thing to note is that much of the curriculum of information competence programs was not thought, as recently as a generation ago, to be the province of a university. First year students were assumed to know how to use a catalogue, how to locate articles in indexing and abstracting services, how to evaluate resources and cite their sources in papers. A student admitted to a university was assumed to have acquired library skills and critical thinking at a previous level of education. This is clearly no longer so and, as with remedial education in all subjects, the library's information competence programme has to deal with remedying past lacks as well as imparting university level knowledge and skills.

Envoi

Here is the fundamental challenge. We need to engage in a range of assessment that is outward looking—that is aligned with the goals of the university rather than internally focused. Those assessments should concern the educational attainments of our students, the research and teaching attainments of our professors, and the service our institutions render to society. However, to a great degree we lack measures that fall between the Scylla of bean counting and the Charybdis of subjective measures of perception. It is incumbent on us, therefore, to develop new nuanced measures that address the

value of the library to the university—measures that deal with real educational and societal outcomes and that demonstrate to all that which we know in our hearts—that a truly educated person is one who can interact autonomously with the human record with confidence and skill.

Thank you.

ⁱ. “*The louder be talked of his virtue/ The faster we counted the spoons.*” Emerson, Ralph Waldo. *Conduct of life, Worship*. Boston: Ticknor & Fields, 1860.

ⁱⁱ. Dickens, Charles. *Hard times*. 1852.

ⁱⁱⁱ. Pritchard, Sarah M. Determining quality in academic libraries. *Library trends*. v.44 (winter 1996) p.591

-
- iv. Lancaster, F. W. The measurement and evaluation of library services. [1st ed.]. Washington, DC : Information Resources Press, c1977. (Second ed. by Sharon Baker and Lancaster, 1991.)
- v. NCLB: No Child Left Behind Act of 2001. PL 107-110, Jan. 8, 2002, 115 Stat. 1425
- vi. SBX1 2: Education Accountability: High School Exit Examination. O'Connell. Chapter 1, California Statutes 1999. (First Extraordinary Session).
- vii. Lindauer, Bonnie Gratch. Defining and measuring the library's impact on campus-wide outcomes. *College & research libraries*. November 1998. p. 549.
- viii. McClure, Charles and Cynthia L. Lopata. Assessing the academic networked environment. Washington, DC: Coalition for Networked Information, 1996. p.5.
- ix. Pritchard, Sarah M. Determining quality in academic libraries. *Library trends* v.44, no. 3 (1996) pp. 572-594. This article is an interesting document that, among other things, is a product of the transition between and overlap of management theories and outcomes assessment.
- x. Powell, Ronald R. Impact assessment of university libraries. *Library & information science research*. V.14 (192 p. 254. *Quoted in* Lindauer. *Op.cit.* p.550.
- xi. Murphy, Dermot F. Developing a culture of evaluation. www.tesol-france.org/articles/murphy.pdf
- xii. Wallace, Danny and Connie Van Fleet, eds. Library evaluation. Englewood, Colo.: Libraries Unlimited, 2001. pp. 3-4.
- xiii. Hernon, Peter and John R. Whitman. Delivering satisfaction and service quality. Chicago: ALA, 2001.
- xiv. *Ibid.* pp. 4-5.
- xv. See a summary of some studies by ALA's Washington Office
<http://www.ala.org/ala/washoff/lba/EconBenesLibrys1pg.pdf>
- xvi. Missingham, Roxanne. Libraries and economic value: a review of recent studies. *Performance measurement and metrics* vol.6, issue 3 (2005)
- xvii. See a list prepared by the American Association of School Libraries at:
<http://www.ala.org/ala/aasl/aaslproftools/resourceguides/studentachievement.htm>
- xviii. Middle States Association of Colleges and Schools. Guidelines for librarian evaluators. Philadelphia, 1997. *Quoted in* Lindauer. *Op. cit.* p. 547
- xix. www.libqual.org
- xx. Wallace, Danny and Connie Van Fleet, eds. *Op.cit*
- xxi. Gorman, G.E. and Ruth Miller. Collection management for the 21st century: a handbook for librarians. Westport, Con.: Greenwood, 1997.
- xxii. Norlin, E. Reference evaluation: a three-step approach--surveys, unobtrusive observations, and focus groups at the University of Arizona. *College & research libraries* 61(6): Nov. 2000. 546-53.
- xxiii. Drabenstott, Karen, Schelle Simcox, and Martie Williams. Do librarians understand the subject headings in library catalogs?(report on the extensive study underway to determine if users and librarians understand the subject headings in catalogs) *Reference & user services quarterly* 38.4 (Summer 1999): p369.
- xxiv. Nitecki, Danuta. SERVQUAL: measuring service quality in academic libraries. *ARL newsletter*, 1997.
<http://www.arl.org/newsltr/191/servqual.html>
- xxv. Parasuraman, A., Leonard Berry, and Valerie Zeithaml. SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. *Journal of retailing*, 1988, vol. 64, no. 1, pp. 12-40.
- xxvi. See, for example, the statement from Oregon State University's library
http://osulibrary.oregonstate.edu/instruction/w121/outcomes_all.html
- xxvii. http://osulibrary.oregonstate.edu/instruction/ug_comp.html
- xxviii. Here explicated by the CSU Los Angeles library College Services and Instruction Task Force
<http://www.calstatela.edu/library/outcomes.htm>