

Counting, Counting, and More Counting... Let's Begin the Countdown to Counting What Really Counts

James Rettig

The Genesis

The catalyst to write this paper was "Assessing the Value of Academic Libraries: Strategies, Tools and Techniques," an excellent one-day workshop sponsored by Towson University and the Maryland ACRL chapter March 21, 2012. Dr. Meagan Oakleaf of the Syracuse University iSchool led the workshop, basing it on her *The Value of Academic Libraries* report, issued by ACRL in September 2010.¹

Through the course of the day Dr. Oakleaf discussed various sources of data an academic library can draw on to demonstrate its contribution to student learning, faculty research and teaching, and fulfilling the institution's mission. The report identifies "surrogates for library impact on student success," among them "internship success, job placement, job salaries, professional/graduate school acceptance, and marketable skills."² The report also counsels that libraries can "partner with campus colleagues in order to leverage existing data sources, including internship evaluation reports, career service records, alumni surveys, and records of individual students' library behaviors."³ And it suggests various correlations a library can draw between these and types of data about the library and its services. Such correlations can support, but rarely prove, the library's case that its work makes a positive difference in student learning and faculty research and teaching.

In a day focused on the importance of outcomes assessment, I thought about another, not unrelated ACRL program, the annual ACRL Trends and Statistics survey which I and some other Towson workshop participants were responsible for submitting the next month.

Among its suggested data sources for drawing the correlations mentioned above, ACRL's *Value of Academic Libraries* report cites circulation counts, online resource and Web site use statistics, space use data, gate counts, ILL use, instructional session enrollment, presence of the library in the campus course management system, library expenditures, and tutorial logins. Some of these data points are included in the annual ACRL questionnaire. Data such as instruction session enrollment and library expenditures are inputs. Indeed, most of the questions in the 2010-11 survey asked for inputs, as does the 2011-12 survey. As a library administrator who has been responsible for the response to this survey since 1998 when it expanded to all four-year colleges and community colleges, I have increasingly had questions in my mind about this survey and the biennial Academic Libraries Survey conducted by the federal government's National Center for Education Statistics (NCES).⁴ One question has been why two very similar surveys are needed, especially when in some cases they define the same data point in two different ways.

James Rettig is Library Director and Associate Dean for Information Services, Nimitz Library, United States Naval Academy, e-mail: rettig@usna.edu

With the exception of an appended annual set of questions about trends, the questions in the ACRL survey are identical to the questions in the “ARL Statistics” survey, the Association of Research Libraries’ flagship annual data collection effort. That effort is guided by the advice and oversight of the ARL Statistics and Assessment Committee. ACRL uses the ARL survey with ARL’s permission.⁵ The ARL Statistics data are also made available to non-member libraries through an Interactive Edition of the ARL Statistics that has online analytics features, a pioneering effort of publishing data on the Web dating back to 1994-95 during the earlier days of the Mosaic browser.⁶

ACRL’s additional trends questions are developed by ACRL’s Academic Library Trends and Statistics Survey Editorial Board. With that exception it is an ARL survey; but when deployed by ACRL, it prominently has ACRL’s name on it. Most libraries who receive the survey from ACRL see ACRL’s name in emails and on associated documents. Because it is best known to its non-ARL recipients as the “ACRL Statistical Questionnaire” and because they receive it from ACRL, my references to the survey (as the ACRL survey) follow the example and precedent ACRL has established in those emails and survey documents.⁷

Some of that survey’s questions have seemed odd or confusing to me. During that March 2012 workshop I wondered why, for example, should libraries report annual expenditures for “contract binding” when the Binding Industries Association exists? What is the value to libraries of collecting and reporting this data point? Volume count has been a prominent fixture in one form or another of academic library statistical surveys since 1908 when James T. Gerould of the University of Minnesota “issued a brief unpublished report containing data on five variables in fourteen state university libraries.”⁸ The 2010-11 ACRL survey included a series of questions on “volumes in library.” Response to the question on “monographic volumes purchased” instructed respondents to “include e-books” with the added instruction to “provide a footnote explaining how many e-books you are reporting, preferably by specifying the products and the number of titles.” The subsequent question asked if the “basis of volume count” was a physical count or a bibliographic count with the explanation that “A physical count is a piece count; a bibliographic count is a catalog record count.” How can e-books be included in a count of physical pieces?

Issues

This volume count issue has a long history.

At its 24th annual meeting, ARL accepted a proposal by Robert Downs that the members should count their holdings in bibliographic units rather than physical volumes. Two years later, in 1947, a survey of members showed that half were counting in bibliographic units and half in physical volumes. In 1948, ARL appointed a new Committee on Counting Library Holdings, chaired by Guy Lyle. In 1949, at the 33rd annual meeting, the Lyle committee recommended that holdings be counted in physical volumes, rather than bibliographic units. Speaking to the motion, Downs said ‘that he had once thought uniformity possible but that he had become disillusioned on this subject and believed that no action taken here would have much effect.’ The members thereupon voted to record in the annual statistics whether counts were in physical volumes or bibliographic units; but no further action was taken on the Lyle proposal.⁹

The 2011-12 survey includes e-books in each library’s total volume count, but has eliminated ambiguity by using only the bibliographic count metric for e-books. It is a welcome step towards clarity. But questions about the value and significance of volume counts abide.

ACRL and NCES take different approaches to the perennial volume count issue. ACRL calls for a “title” count, complemented by a separate volume count. It bases the title count on the ANSI/NISO Z39.7-2004 standard. That standard folds serials titles into the overall title count by noting that

A book or serial title may be distinguished from other such titles by its unique International Standards Book Number (ISBN) or International Standard Serial Number (ISSN). The definition applies equally to print, audiovisual, and other library materials...Two subscriptions to Science magazine, for example, are counted as one title.¹⁰

NCES’s FY2012 materials count—which “does not cover all materials”—uses the ANSI/NISO Z39.7-1995 standard. Both the ACRL and NCES counts define

volume as “a single physical unit of any printed, type-written, handwritten, mimeographed, or processed work, distinguished from other units by a separate binding, encasement, portfolio, or other clear distinction which has been cataloged, classified, and made ready for use, and which is typically the unit used to charge circulation transactions.”¹¹ Both include “duplicates and bound volumes of periodicals;” however NCES also includes “print photographs,” a document type on which ACRL is silent—unless, that is, they are included among “audiovisual materials (cartographic, graphic, audio, film and video, etc.)”¹² Both include “Government document volumes that are accessible through the library’s catalog regardless of whether they are separately shelved.”¹³ The ACRL survey for 2011-12 data adds the complicated caveat that

Documents should, to the extent possible, be counted as if they were in bound volumes (e.g., 12 issues of an annual serial would be one or two volumes. Title and piece counts should not be considered the same as volume counts. If a volume count has not been kept, it may be estimated though sampling a representative group of title records and determining the corresponding number of volumes, then extrapolating to the rest of the collection

and applying any of three formulas (two of them based on linear footage of shelving occupied by documents and the third on volume equivalency).¹⁴ This brings to mind IRS 1040 instructions, but without the helpful worksheet to structure the calculations.

The NCES and ACRL surveys will almost certainly yield two different volume counts for any library that applies both precisely. Given the definitional differences related to volume count—undoubtedly the most revered library metric, and perhaps equally doubted—one wonders how ACRL and NCES data can be used without reservation for comparative purposes.

The overriding question I have had for more than a decade is why two surveys—why both the ACRL survey and the NCES survey? There is a historic reason for ACRL’s decision in the late 1990s to expand its previously selective survey to all academic libraries in the United States and Canada. One was to provide academic library statistical data in a more timely way than was available from biennial federal government

surveys which suffered from long lags between data collection and publication. Conducting the survey annually as well as using the Web to collect data responded to academic librarians’ “need for more current and comprehensive data about other libraries than was easily accessible.”¹⁵

Nothing illustrated my questioning of the need for both the NCES and ACRL surveys than the vexed question of serials subscriptions, complicated ever since the advent of bundled e-journal packages and the presence of titles in multiple online aggregations. In 2008 ARL changed its definition for counting serials from subscriptions to titles so that “If a title appears in both print and electronic form and a library has acquired it through several different providers, it would be counted as a single, electronic, purchased title.”¹⁶ In this spirit the ACRL instructions for 2010-11 stated that “If a purchased [serial] title includes electronic access to the title, count that title ONLY ONCE (DEDUPED) as electronic only. If a database includes full-text and abstracted titles, the number of full-text titles can be counted.”¹⁷ “Can?” Does this mean that each reporting library is free to decide for itself whether or not to include full-text serials included in a database? If so, what does that mean for the value of this data point for comparisons of libraries? In contrast, the NCES instructions for counting serials for FY2010 stated “Report the total number of titles in all formats. If the title comes in both paper and electronic form, count it twice. Count each individual title if it is received as part of a publisher’s package.”¹⁸ To further complicate serials counting, NCES recorded “indexing and abstracting services that may contain full-text” separately.¹⁹ Why such a dramatic difference between ACRL and NCES?

Then there is the issue of survey responses from institutions that include a main library and its branches and other substantial libraries such as a medical school library or law school library—or both. NCES notes that “Branch and independent libraries are administered either by the central library or, as in the case of some libraries (such as law, medical, etc.), through the administrative structure of other units within the university.”²⁰ The 2010-11 ACRL definition is similar, but different: “A branch library is administered *either* by the central library *or* (as in the case of some law and medical libraries) through the administrative structure of other units within the university.” ACRL states a preference that a survey response *not*

include data for branch campuses, but is ambiguous about whether or not a single campus should consolidate data for its main library and branches. Applying this place-based principle, NCES in 2010 was explicit: “Include data for all branch and independent libraries on the campus.”²¹ The NCES FY2012 instructions repeat this statement. If every university had the same academic structure of schools and libraries, such consolidated library data could be used for comparisons. But structures vary widely, greatly limiting the value of these surveys for libraries that want to compare themselves to comparable libraries that are part of multi-library systems. ARL has addressed this problem, at least for ARL libraries, by issuing separate reports for its member institutions’ academic health sciences libraries and academic law libraries.²²

Again, why two surveys—why both ACRL and NCES? Couldn’t their sponsor merge their surveys? And more importantly, why two surveys that count some of the same things in two different ways and yield different quantitative results?

Welcome Changes

Change is afoot—change for the better. NCES data has a varied history in the Integrated Postsecondary Education Data System (IPEDS). “NCES surveyed academic libraries on a 3-year cycle between 1966 and 1988. Between 1988 and 1998, the ALS was collected on a 2-year cycle. Beginning with fiscal year 2000, the ALS is no longer a component of IPEDS, but remains on a two-year cycle.”²³ However, the FY12 NCES Academic Library Survey (ALS) is intended to be the last. “Starting in 2013-14, IPEDS will begin to collect ALS data elements, on an annual basis.”²⁴ Some of the data requested in the biennial ALS has been reported annually at the institution level. Change has also been motivated by the intent to decrease the reporting burden and “retain federal data necessary for policy making and analysis, while also improving response rates.”²⁵ Since institutions are required to respond to IPEDS surveys, the response rate will certainly improve. A Technical Review Panel has made recommendations about which data elements should be reintegrated into IPEDS. These deal with expenditures for library staff, collections, circulation, interlibrary loans, and virtual reference (a yes/no question). Circulation and ILL are included because the panel concluded that “some measure of usage must be integrated into IPEDS” and “agreed that the amount of circulation transac-

tions is an indicator of usage as well as an important library service.”²⁶ Assuming for simplicity’s sake that the Technical Review Panel’s recommendations are accepted as proposed, questions remain.

Circulation data have been gathered since at least 1849 when Charles Coffin Jewett included it among eighteen questions in a survey “of all ‘public’ libraries in the country.”²⁷ Their value cannot go unquestioned. Does an individual library’s report of its circulation in a given year include renewals? The length of loan periods for different categories of borrowers is a variable that needs to be accounted for if circulation data is used in comparing libraries. Some libraries have a uniform end-of academic term loan period, sometimes applicable to all categories of borrowers, other times applicable only to select categories. Some require materials to be returned in person after a certain number of renewals; they are then checked in and immediately checked out to initiate a new loan and possible subsequent renewal(s). Luzius has pointed out that “Each manufacturer’s circulation software has inherent ways of counting circulating items that can play a part in libraries’ final statistics.”²⁸ More significantly “The number of materials that are borrowed is a poor surrogate for expressing their value to those who do the borrowing, and for capturing a sense of their impact.”²⁹ It is probably a rare library user—I know I am not such a user—who has borrowed a book with the best of intentions but has returned it without having read any of it. Nevertheless IPEDS may enshrine circulation as “an indicator of usage as well as an important library service.” It is surely the latter, but not necessarily the former.

It remains to be seen how each of the collections data elements to be included in IPEDS (books, serials, databases, media) will be defined. What is clear is that physical collection resources and digital/electronic will be treated separately in response to the panel’s concern “that the ALS does not make a distinction between a physical count of materials and an electronic/digital count of materials. Thus, the panel suggested capturing the allocation of online and physical materials to allow institutions to make peer comparison of the distribution of resources.”³⁰ This is in line with ARL’s shift in focus from things held to investments made.

ARL has made welcome changes in its 2011-12 statistical survey of its member libraries. Expenditure for one-time resource purchases (monographs, seri-

als, other materials) have been consolidated into a single response. Electronic books are included in the overall volume count, but the physical count/bibliographic count distinction clearly applies only to the print volume count. Library use is included for database searches and federated database searches, following COUNTER definitions. The ARL survey does not wade into the vexed issue how to count serials subscriptions. It no longer asks that question. Because it uses the ARL survey, these changes also appear in the 2011-12 ACRL survey.

Quality Issues

In his wide ranging history of numerous efforts to create a national, ongoing statistical portrait of American libraries, Williams notes that “Three problems have to be addressed in considering the quality of data that have been collected about libraries in the past 150 years: completeness (or comprehensiveness), reliability, and validity.”³¹ The quest for completeness, reliability, and validity has been elusive, but not without progress, not just in the United States, but also in other countries and regions of the world. Often these three criteria intersect, making it a challenge to fulfill all three.

Completeness

Clear definitions accepted by libraries contributing data to a survey lay the foundation to completeness, reliability, and validity, especially for national or multi-nation library statistical programs. Young has noted that “Any use of library statistics depends upon a healthy and consistent data gathering and definitional component which has provided time-series of comparable data to cumulate, compare, and aggregate through interpretative processes which allow conclusions to be drawn.”³² An effort to harmonize annually reported data in Denmark, Finland, Iceland, Norway, and Sweden reported in 1990 that it encountered problems over “the type and number of libraries to be included in each library category...categorization of staff, since their educational backgrounds differ from country to country...[and]consensus on the exact definition of these Categories.”³³ Similarly, a project to develop “an online statistics service for a group of 22 academic libraries in Hong Kong, Malaysia, Singapore, and Thailand” fell short of its ambitions because participants perceived “insufficient or unclear definition of some data elements; reluctance to make some

data public; [and] lack of a critical mass of data.”³⁴ A benchmarking project for Dutch academic libraries begun in 1998 and implemented in 2000 fell short in reliability because “Large differences between similar libraries or between consecutive years may indicate that the data is incorrect.”³⁵

An effort to collect global library statistics, the subject of a conference held in conjunction with the 2008 IFLA conference, reported having reached “consensus between the cooperating groups of IFLA, ISO and UNESCO on a list of relevant library statistics... based on definitions in the international standard ISO 2789.”³⁶ However completeness was a bigger challenge because a test “in Latin America and the Caribbean demonstrated that many data in the questionnaire **could not yet be collected** by the countries but at the same time proved the feasibility of the project by the respectable percentage of data that **could be collected.**”³⁷ [emphasis in original] These examples validate Williams’s statement that “It is a given that all library data compilations are incomplete.”³⁸

Another impediment to completeness is worry about how data will be used. A statistical survey of public libraries in South Africa circa early 21st century received many incomplete questionnaires. Discussions during a workshop in Cape Town in November 2002 revealed “that some librarians are reluctant to collect statistics for fear they might be used to the disadvantage of their library.”³⁹ For example, “an unintended consequence of the potential for comparison among [library] systems might be that a library whose spending is lower than another might be held up as efficient, rather than under-resourced.”⁴⁰ In a personal phone conversation with Ray English of Oberlin College several years ago, Mr. English explained to me that some members’ concern that their institution will look bad compared to others is one reason the Oberlin Group does not share its statistics beyond its members. Such concerns have their history beyond the Oberlin Group.

Stubbs quotes correspondence between Phineas Windsor, librarian at the University of Illinois and his president after Windsor sent President Edmund James Gerould’s library statistics, noting that Illinois had fewer volumes than Oberlin. The president wrote “I note that while Illinois only added 29,000 volumes during the year 1913-14 and Yale added 37,000, we spent for the 29,000 \$86,000, while Yale spent only \$34,000. What accounts for this discrepancy? Is Yale

getting much better bargains, or are we buying more expensive books, or what?”⁴¹ As described by Stubbs, librarian Windsor’s defensive response to President James, running a page and a half, prefigures the footnotes respondents add to the quantitative responses to ACRL survey questions to qualify the reported data and place it in local context.

In 1991 Williams said that “The Gerould/Princeton/ARL statistics are the longest—and arguably the best—continuous series of data in the United States,” a judgment still widely held.⁴² Perhaps, behind their locked door, the Oberlin Group statistics, albeit not as long a series, are of equal quality and value. If so, it is a mournful irony that in a profession that deeply values sharing and dissemination of information, the library directors of eighty “selective, top-ranked liberal arts colleges” impose a permanent embargo on their data that could be useful to other liberal arts colleges. ARL once debated this issue internally. During ARL’s meeting in May 1986 in the directors’ discussion about the ARL statistics, James Govan of the University of North Carolina said “I am really concerned about the principle involved in our calculating and not publishing the Index. We are an organization that is supposed to stand for the free flow of information, to oppose any kind of suppression of information.”⁴³ Public access on the ARL Web site to a variety of statistical reports for a number of years demonstrates ARL’s commitment to transparency and access to information.

Reliability

Regarding reliability, Williams says it “asks whether the measuring instruments gave dependable and consistent answers from one library to another.”⁴⁴ In that discussion at the 1986 ARL meeting, MIT’s Jay Lucker noted that “the ARL data are looked at as fairly stable...not only the best there are, but...we know despite that there is a lot of inconsistency in the data.”⁴⁵ Richard Dougherty of the University of Michigan, addressing “access measures” under consideration for inclusion in the ARL Index said, “Not only must we arrive at a definition we can all accept, we must transmit that information to the various members of the staff who are responsible for collecting the information.”⁴⁶

If the reliability of even the vaunted ARL statistics are subject to doubt, one has to wonder about the reliability of the NCES and ACRL survey data. Dick Daugherty’s 1986 caution that the data’s “import-

tance must be transmitted to those on staff who are responsible for gathering the information” applies to all libraries and their directors.⁴⁷ Based on my experience of becoming the library director at two liberal arts institutions, local practice in data gathering can trump standards and definitions used in national surveys. The ways in which individual staff, especially long serving staff, gather data and thereby implicitly define data elements don’t always translate accurately into ACRL or NCES data elements. Young observed that “Noncompliance with the standard results from ignorance, inertia, and individualism.”⁴⁸

This phenomenon is probably most pronounced in tracking reference statistics. The literature on reference statistics, especially the distinction between directional and reference questions, is voluminous and a subject worthy of extended treatment in itself. It, probably more than any other library metric, illustrates the issue of individualism, sometimes even at the departmental level within a single library, and how individualism can undermine reliability of reference transaction data. For example, a self-study at the University of New Mexico in 2005 made it “apparent that the ways different branches of the library across the campus gathered reference statistics were not consistent and not always accurate.”⁴⁹

Validity

Williams wrote that “The heart of the problems about data quality is validity, which asks if you are measuring what you want to measure.”⁵⁰ It is a question at least as old as Charles Ami Cutter’s 1889 ALA presidential address. It is a question we librarians have increasingly been asking ourselves in the past two decades. *Measuring Academic Library Performance* illustrates the movement away from input measure to output measures.⁵¹ Yet input measures remain the mainstay of the national academic library surveys. Regional accrediting agencies have rightly pressured colleges and universities to develop assessment programs. This has been challenging for faculty who have extended contact with students and who have ready access to measures of student success in their courses. Grades do not in and of themselves constitute an outcomes assessment measure of student learning, but they come closer to filling that role than academic libraries’ abundant input data and their output data.

At the 1995 ACRL conference Sarah Pritchard raised questions about benchmarking, an exercise

dependent upon comparative institutional data. She concluded that

we may be trying to quantify the unquantifiable, that is to measure the production and transmission of recoded knowledge and its impact, using techniques that are inherently linear and atomistic (e.g., having discrete data points on annual time lines) to describe a process that is neither. That process is increasingly non-linear and riddled with externalities, and neither the services nor the products can be consistently or completely broken down into controllable units that are the same for every institution.⁵²

I echo Pritchard's question from nearly twenty years ago: "Do we really know why we are collecting much of the data we tally in our libraries?"⁵³ It is a question worthy of repetition. Clearly targeting the way in which libraries used the ARL statistics year after year long after Illinois' president James quizzed librarian Windsor on the meaning of comparative library statistics, Herb White asked "Do we know how many books we *really* have? Should we care? Should anyone care?"⁵⁴ [emphasis in original] He also anticipated Pritchard's question with the exhortation to "Stop counting and reporting the things that aren't worth counting."⁵⁵

ARL and ACRL and librarians are seeking answers. The elimination of data elements such as expenditures on binding, the shift from counting types of materials to reporting total investment in access to information, transcending the quagmire of how to count serial subscriptions—these are steps in the right direction. ARL-supported initiatives such as LibQUAL+, DigiQUAL, and its biennial assessment conference are helping libraries gather and analyze data that largely lies outside the scope of the traditional academic libraries surveys. Programs, reports, and toolkits under the umbrella of ACRL's "The Value of Academic and Research Libraries" initiative are doing the same.⁵⁶

Yet those input-dependent statistical surveys are still with us. Williams notes that

Charles Ami Cutter was so dissatisfied with the results he was getting from the statistical reports that he said in his ALA presidential address in 1889: "I cannot help regretting the amount

of time that was wasted on statistics. They are interesting, but they are costly to prepare and print, and I would rather see the time spent on making the library more useful."⁵⁷

In the paper that was the catalyst for the directors' discussion of the ARL at their 1986 meeting, Kendon Stubbs of the University of Virginia declared that "The Statistics Committee believes that the ARL statistics can be redeemed."⁵⁸

RECOMMENDATIONS

In the spirit of Cutter's pursuit of the goal to make the library more useful, I believe that the national statistical surveys can be redeemed and can help demonstrate how the library is useful—and even play a role in making it more useful. I have several recommendations for action, most calling for research that is doable.

We should count only what counts.

ARL—ideally in partnership with ACRL's Academic Library Trends and Statistics Survey Editorial Board, as well as in consultation with IPEDS—can lead the way in making the surveys more useful. Through recent changes in the data elements included in its survey, ARL has demonstrated willingness to examine these elements critically. Its Statistics and Assessment Committee can and should continue that process. As a part of the examination, it should also conduct a study of how ARL libraries use each data element. These uses include internal management decision making, benchmarking and for comparisons to other research libraries, and assessment. Those data elements honored more by being collected than by being usefully used are dispensable and should be discontinued. If the (in)famous titles/volumes held measure serves no more purpose than it did in Phineas Windsor's day, the answer to White's question is that nobody should care any longer how many books or other information sources a library has. The serials issue may be the harbinger for this. Does it matter to a student how many books a library owns if it doesn't have a book the student needs?

ARL could also publish case studies of how the various data elements are used by ARL libraries, both for internal uses and for comparisons.⁵⁹ More importantly, those case studies could explain how those uses of the data contribute to local assessment efforts—in

other words, how they make the library more useful. The ARL Profiles program offers a possible model. In 2009-10 ARL's Research Library Leadership Fellows identified themes that merit enhanced reporting among member libraries and reported on members' responses to a dozen suggested themes. These included digital publishing, collaborations, and social networking and mobile applications.⁶⁰ The abiding value of this effort is that it established a process that can be iteratively and constructively applied to refining the statistical surveys.

We should learn how library staff actually track, record, and report data asked for in the surveys.

As the University of New Mexico self-study demonstrated and ARL directors' comments made in 1986 reinforce, there is good reason to question the reliability and validity of the data reported. As one library administrator told me, "We try not to make it up." In recent years we have learned the value of studying the information seeking and learning behaviors of our students, especially undergraduates. In the work that produced *Studying Students: The Undergraduate Research Project at the University of Rochester* Nancy Foster and Susan Gibbons demonstrated the value of applying ethnographic studies to understand habits, perceptions, and preferences of those we serve.⁶¹ In *College Libraries and Student Culture: What We Now Know* Lynda Duke and Andrew Asher have built on Rochester's groundbreaking work and reported findings from studies of students at five universities in Illinois.⁶²

What works with students should work with and for us. We need a study of a cross-section of academic libraries of varied type and size to learn just how the library staff responsible for the survey's data produce and report it. How well do they know the definitions of the data points? How well do they adhere to those definitions in the data they report?

We know there are year-to-year variations in data reported from a single institution. Such a study can help us understand how such variations come about. Are they accurate reports of data or are they the result of changes in local procedures, policies, or personnel? When there is an anomaly from one year to the next, the online form for the ACRL survey alerts the reporter to this and prompts the reporter to explain it in a note. How great are those quantitative anomalies? What do they signify?

We should study the notes that explain or qualify individual libraries' responses.

The notes to the ACRL survey are themselves worthy of a research study. Such a study may reveal confusion about interpretations of definitions, if there are data elements with which a consequential number of libraries have difficulty or for which they are unable to supply data, and could uncover patterns or pockets of elements for which data must be judged inconclusive or unreliable.

Ranganathan applied to Academic Library Statistics

Many libraries use survey data for benchmarking and comparisons with libraries at peer or aspirant institutions. The very affordable ACRLMetrics product provides "access to ACRL and NCES academic library statistics (2000 to present) plus...a select subset of IPEDS data specific to academic libraries."⁶³ This tool allows users to read the notes. Many use ACRLMetrics to create groups of similar libraries to generate comparative data and benchmarking information. This tool makes it easy to calculate various metrics explained in Dugan, Herson, and Nitecki's *Viewing Library Metrics from Different Perspectives*.⁶⁴

If, as they do, libraries are going to calculate ratios and draw correlations between those ratios and local library quality as well as relative library quality compared to other institutions, it is vital that we collectively know the reliability and validity of the statistical data that those correlations and comparisons rest on. Until research verifies or demonstrates the limits of their reliability and validity, metrics, ratios, and comparisons rest on faith. Is that faith well founded? We really don't know. We do know there is reason to question their reliability and validity.

Two of Ranganathan's timeless laws of library science are:

- Save the time of the reader.
- A library is a growing organism.

One thing that could contribute to increased reliability of the data reported on the surveys would be accompanying tools to save the respondents' time and to assure understanding of definitions. Three PDF documents support the 2011-12 ACRL survey: Instructions, including definitions, for completing the survey; an FAQs document; and a worksheet.⁶⁵ When a library actually enters the data online, some simple addition calculations are performed. As long as the ACRL survey's instructions on how to count titles and volumes

are as complicated as they are today, there is considerable room for error and inconsistency, both intra-institutionally year-to-year and inter-institutionally.

There is an example to emulate. Anyone who has used TurboTax™ or a similar product to calculate their personal federal and state income taxes, knows how well these tools integrate instructions, definitions, calculators, and even IRS publications. The ACRL survey would benefit greatly from a comparable step-by-step tool, updated annually. Just as ACRLMetrics simplifies analysis of reported data, a TurboTax-like tool can provide similar convenience at the input stage and concomitantly enhance data reliability. Such a tool could also aggregate data for a “main” library and campus branches, but when data is uploaded from the tool, also disaggregate it for each major library on a campus and report it both ways. This would allow users, within ACRLMetrics (or other means) to use the aggregated or disaggregated institutional data set more suitable to their purposes, especially when making inter-institutional comparisons.

We all have a stake in improved national library statistical surveys. As NCES migrates into IPEDS and collects highly selective library data, the ARL survey and its ACRL-branded clone will become more important. Surely ARL and ACRL can collaborate to launch an effort, more nimbly than IPEDS, to revise their shared survey to count what counts.

I believe that the ARL statistical survey (and thereby in ACRL’s identical survey) can be redeemed. It will be redeemed when research demonstrates that we can believe the data, or when we make the changes research reveals must be made to make the data believable. As we identify the statistical metrics that will meaningfully support our assessment efforts and drop meaningless or insignificant measures and add meaningful measures, and know that data reported has high reliability and validity, library statistical surveys will fulfill their potential to help make the library more useful.

Addendum

The author thanks Dr. Martha Kyrillidou, Senior Director, ARL Statistics and Service Quality Programs, at the Association of Research Libraries for her comments on early drafts of this paper and for generously sharing her unequalled knowledge of the ARL Statistics and the processes ARL uses to review and renew this resource.

Notes

1. Meagan Oakleaf, *The Value of Academic Libraries: A Comprehensive Research Review and Report* (Chicago: Association of College and Research Libraries, 2010). http://www.ala.org/acrl/sites/ala.org.acrl/files/content/issues/value_val_report.pdf
2. *Ibid.*, 109.
3. *Ibid.*
4. National Center for Educational Statistics, “Academic Libraries Survey Questionnaires and Instructions.” http://nces.ed.gov/surveys/libraries/aca_questdefs.asp
5. Association of Research Libraries, *ARL Statistics* (annual). Available in pdf through <http://publications.arl.org/> and available as the Interactive Analytics ARL Statistics edition through <http://interactive.arlstatistics.org/> to non-member libraries.
6. Martha Kyrillidou, “To Describe and Measure the Performance of North American Research Libraries.” *IFLA Journal* 27, no. 4 (2001): 257-263.
7. For example, two of the three explanatory documents for the 2011-12 survey distributed by ACRL are titled, respectively, “ACRL Statistics 2011-12 Worksheet” and “ACRL Statistics Questionnaire 2011-12: Instructions for Completing the Questionnaire.” The footer on most pages of both of these documents carries the statement “The Association of Research Libraries (ARL) has made the ARL Statistics survey available under a Creative Commons Attribution-NoDerivs—NonCommercial 1.0.” However the footer on the two pages of the worksheet document devoted to the trends questions states: “The 2012 Trends questionnaire was developed by the 2013 ACRL Academic Library Trends and Statistics Survey Editorial Board.” The attribution in the footers is the only mention of ARL in the worksheet document; there are several references to ARL definitions in the instructions document.
8. Robert V. Williams, “The Making of Statistics of National Scope on American Libraries, 1836-1986: Purpose, Problems, and Issues,” *Libraries and Culture* 26 (Spring 1991): 467.
9. Kendon Stubbs, “Lies, Damned Lies, . . . and ARL Statistics?” in *Research Libraries: Measurement, Management, Marketing: Minutes of the 108th Meeting* (Washington: Association of Research Libraries, 1986), 84-85.
10. “ACRL Statistics Questionnaire, 2011-12: Instructions for Completing the Questionnaire,” 2. <https://acrl.countingopinions.com/docs/acrl/12acrlinstruct.pdf>
11. *Ibid.*, 3.
12. *Ibid.*, 2.
13. *Ibid.*, 3.

14. Ibid.
15. Hugh Thompson, "ACRL Wraps Up Year 1 of Academic Library Statistics Project: How You Can Get the Results," *College and Research Libraries News* 60 (October 1999): 760.
16. Martha Kyrillidou, "ARL Statistics: Redefining Serials Counts and Remaining Relevant in the 21st Century," *Research Library Issues* 262 (February 2009): 19. See also Martha Kyrillidou, "Reshaping ARL Statistics to Capture the New Environment," *ARL: A Bimonthly Report*, no. 256 (February 2008): 9–11, <http://www.arl.org/bm~doc/arl-br-256-stats.pdf>; and Martha Kyrillidou, "The Impact of Electronic Publishing on Tracking Research Library Investments in Serials," *ARL: A Bimonthly Report*, no. 249 (December 2006): 6–7, <http://www.arl.org/bm~doc/arlbr249serials.pdf>.
17. For background on ARL's work on how to account for serials, see minutes of the ARL Statistics and Assessment Committee at <http://www.arlstatistics.org/about/committee/committee>.
18. National Center for Educational Statistics, "Instructions for the Academic Libraries Survey—FY2010," n.p. http://nces.ed.gov/surveys/libraries/pdf/ACQuestionnaire_FY10.pdf
19. Ibid.
20. Ibid.
21. Ibid.
22. *ARL Academic Health Sciences Library Statistics, 2009-2010*. (Washington: Association of Research Libraries), 2012. *ARL Academic Law Library Statistics, 2009-2010*. (Washington: Association of Research Libraries), 2012.
23. Report and Suggestions from IPEDS Technical Review Panel #35, Reintegrating the Academic Libraries Survey into IPEDS, 1. https://edsurveys.rti.org/IPEDS_TRP/documents/TRP35_SummaryPackage_Suggestions_final.pdf
24. Ibid.
25. Ibid., 4-5.
26. Ibid., 8-9.
27. Williams, 465.
28. Jeff Luzius, "A Look at Circulation Statistics," *Journal of Access Services* 2 (2004): 17.
29. Alvin M. Schrader, "\$00 Million Circs, 40 Million Reference Questions: What Does This Mean and Does Anybody Care? Getting Beyond Library Statistics to Library Value with Help from Canada's National Core Library Statistics Program," *Argus* 35 (2006): 18.
30. Ibid., 8.
31. Williams, 478.
32. Peter R. Young, "U.S. Library Statistics," 3 *Library Administration and Management* (1989): 174.
33. Ivar A. L. Hoel, "Harmonization of Statistical Data from Nordic Research Libraries," *Libri: International Journal Of Libraries & Information Services* 40 (1990): 217.
34. Cathie Jilovsky, "Singing in Harmony: Statistical Benchmarking for Academic Libraries," 32 *Library Management* (2011): 48, 56, 58.
35. Henk Voorbij, "Ten Years of Experience with Benchmarking in Dutch Academic Libraries," in *Library Statistics for the Twenty-First Century World: Proceedings of the Conference Held in Montréal on 18-19 August 2008 Reporting on the Global Library Statistics Project* (Munich: K. G. Saur): 247.
36. Roswitha Poll, "Where Shall We Go From Here?" in *Library Statistics for the Twenty-First Century World: Proceedings of the Conference Held in Montréal on 18-19 August 2008 Reporting on the Global Library Statistics Project* (Munich: K. G. Saur): 290.
37. Ibid.
38. Williams, 479.
39. Karin de Jager and Mary Nassimbeni, "Towards Measuring the Performance of Public Libraries in South Africa," *South African Journal of Libraries and Information Science* 71 (2005): 41.
40. Ibid.
41. Kendon Stubbs, "Apple, Oranges and ARL Statistics," *Journal of Academic Librarianship* 14 (1998): 231.
42. Williams, 467.
43. "Reports from Discussion Groups," in *Research Libraries: Measurement, Management, Marketing: Minutes of the 108th Meeting* (Washington: Association of Research Libraries, 1986): 33.
44. Williams, 479.
45. "Reports from Discussion Groups," 36.
46. Ibid.
47. Ibid., 37.
48. Young, 174.
49. Carolyn Caizzi, "By the Numbers, Gathering and Using Statistics," *VRA Bulletin* 37 (2010): 19.
50. Williams, 480.
51. Nancy A. Van House, Beth T. Weil, Charles R. McClure, *Measuring Academic Library Performance: A Practical Approach* (Chicago: American Library Association, 1990).
52. Sarah M. Pritchard, "Library Benchmarking: Old Wine in New Bottles?" *Journal of Academic Librarianship* 21 (1995): 495.
53. Ibid., 492.
54. Herbert S. White, "Trouble at the OK Corral University Library," *Library Journal* (September 1, 1987): 155.
55. Ibid.
56. Value of Academic and Research Libraries, <http://www.ala.org/acrl/issues/value>
57. Williams, 480.

58. Kendon L. Stubbs, "On the ARL Library Index," in *Research Libraries: Measurement, Management, Marketing: Minutes of the 108th Meeting* (Washington: Association of Research Libraries, 1986): 18.
59. See Association of Research Libraries, "ARL Statistics® Web-cast," (Washington: ARLVideo, 2012). <http://www.youtube.com/watch?v=fKegMJWaONk&list=PL1180A414DC719544&index=3>
60. Robert Fox, Patrick Reakes, Bryan Skib, Ann Snowman, "ARL Profiles: Emerging Themes," October 13, 2010, <http://www.arl.org/bm~doc/profilesrllfproject.pdf>
61. Nancy Fried Foster and Susan Gibbons, *Studying Student: the Undergraduate Research Project at the University of Rochester* (Chicago: Association of College and Research Libraries, 2007).
62. Lynda M. Duke and Andrew D. Asher, *College Libraries and Student Culture: What We Now Know* (Chicago: American Library Association, 2012).
63. ACRLMetrics, <http://www.acrlmetrics.com/>.
64. Robert E. Dugan, Peter Hernon, and Danuta A. Nitecki, *Viewing Library Metrics from Different Perspectives: Inputs, Outputs, and Outcomes* (Santa Barbara: ABC-CLIO, 2009).
65. See endnote #7 for information about two of these three documents.