arl.wp1 6/20/06
1st rev 6/29/06
2nd rev 9/1/06
3rd rev 10/6/06

Some Alternative Quantitative Library Activity

Descriptions/Statistics That Supplement the

ARL Logarithmic Index

Bruce Thompson

Texas A&M University
and
Baylor College of Medicine (Houston)

Bruce   Thompson   is   Distinguished   Professor   of   Educational
Psychology   and   College   Distinguished   Research   Fellow,   and
Distinguished Professor of Library Sciences, Texas A&M University,
and Adjunct Professor of Family and Community Medicine, Baylor
College of Medicine (Houston).

Q. *What quantitative scores have historically been reported by the ARL Statistics and Measurement Program?*

A. Traditionally, the primary quantitative score reported by the ARL Statistics and Measurement Program has been the "ARL Membership Criteria Index" (hereafter, "ARL Index"). As explained within ARL documentation[1],

> The criteria for academic library membership in the Association of Research Libraries are based partly on quantitative data that provide a view of the range of resources deployed among the existing members of the Association. Statistical analysis shows a high degree of homogeneity in respect to five data categories:
>   --volumes held
>   --volumes added, gross
>   --current serials
>   --total library expenditures
>   --total professional plus support staff
> Each year ARL uses the statistical method of principal component analysis to identify the commonalities in the membership. The analysis is conducted on the 35 charter members of ARL and produces coefficients, or weights, for each of the five data categories. When the data for a given library are multiplied by the weights and summed, the result is a "score" for that library. This process of multiplying by weights and summing is carried out for each ARL academic library. The resulting scores comprise what is known as the ARL Membership Criteria Index. The term "score" in this context is not a judgment about the library's quality or performance. "Score" is a term from principal component and factor analysis that refers to the summation of data. The ARL index score in effect aggregates the five measures of size and resources. Each year the current year's data in the five categories are published in The Chronicle of Higher Education, arranged in descending rank order by the ARL Index scores.
> The weights and data categories can also be applied to the data of non-ARL libraries. This technique is one of the tests used to determine potential members of the Association. Candidates for membership are required to have a score on the ARL index scale of at least -1.65 for the most recent four years in order to be considered. This criterion was established to ensure that new members share the essential characteristics of the existing members in regard to the five measures of size. The membership criteria also include other

---

[1]<http://www.arl.org/newsltr/197/criteria.html>.

requirements to ensure the homogeneity of the membership.[2]

Q. *What prior literature exists with respect to the ARL Index?*
A. A variety of articles and reports on the ARL Index have been published through the years, first by Kendon Stubbs, later by Martha Kyrillidou, as well as by persons unaffiliated with ARL (e.g., Weiner, 2005). Appendix A presents a sampling of these articles and reports.

Q. *What does the ARL Index measure?*
A. By its very nature, the ARL Index is a measure of **resources input** into the library (i.e., print collections and staffing).

Q. *Why are other quantitative statistics necessary today?*
A. First, input statistics were never meant to capture all aspects of research library operation. For example, input statistics do not characterize the **output service quality** that libraries provide to users. Thus, at the October, 1999 ARL Membership Meeting, the ARL Statistics and Measurement Committee and the ARL Research Library Leadership and Management Committee initiated the ARL "New Measure Initiative." Among these initiatives have been projects such as:

--LibQUAL+®, a protocol for measuring users' perceptions of library service quality;[3]
--MINES ("Measuring the Impact of Networked Electronic Services"), a protocol for measuring users' purposes when accessing specific digital services;[4] and

---

[2]See <**http://www.arl.org/stats/factor.html**> for more detail on the membership index. The formulas for recent years used to compute the ARL Index can be found at: <**http://www.arl.org/stats/index/indxform.html**>.

[3]LibQUAL+® has now been completed by more than 700,000 library users at more than 700 libraries around the world (e.g., the United States, Canada, Australia, New Zealand, England, France, Ireland, Scotland, the Netherlands, Switzerland, Denmark, Finland, Sweden, Egypt, the United Arab Emirates, and South Africa). Both (a) the development and (b) the use of LibQUAL+® data to improve service quality have been documented in more than 4 dozen articles published in journals such as College and Research Libraries, IFLA Journal, Journal of Academic Librarianship, Journal of Library Administration, Library Administration & Management, Library Quarterly, Library Trends, Performance Measurement and Metrics, and portal.

[4]The Project MINES for Libraries™ protocol has been used at the Ontario Council of University Libraries (OCUL) across 16 libraries. Local institutional implementations have taken place as part of larger indirect cost studies carried out by various

--COUNTER ("Counting Online Usage of Networked Electronic Resources"), a project formally incorporated in England as a not-for-profit company in 2003, which facilitates the development of standards and protocols involving the recording and exchange of online usage data.

Second, the increased emphasis on libraries providing digital content to users also has exponentially impacted the need to consider quantitative statistics in addition to the ARL Index, which does not consider digital content. The ARL statistics on (a) circulation and (b) reference transactions, reported in Figures 1 and 2, reflect the impacts of this movement by users toward greater use of digital content.

Q. *Are there any technical/statistical differences between the analyses reported here versus those employed in computing the ARL Index?*

A. Yes. The ARL Index is <u>not</u> based directly on the 5 variables (e.g., volumes held, volumes added, gross). Instead, the ARL Index is based on the <u>natural logarithmic</u> values of these five datapoints[5]. For illustrative purposes, below are presented 5 numbers of volumes held, and their respective log values.

Volumes Held	log of Volumes Held
3,800,000	15.15051
1,900,000	14.45736
1,398,000	14.15055
1950	7.57558
3	1.09861

The rightmost values are the kinds of data actually analyzed in computing the ARL Index scores. Clearly, the log values are not in the metric in which most librarians think.

There are arguably some good statistical reasons to use log values in computing the ARL Index scores (e.g., log values make nonlinear dynamics more linear, and focus on relative differences versus absolute differences in data such as collection size). Nevertheless, an alternative focus on data in its more familiar form is not unreasonable, and also is in keeping with the purpose of the present work.

---

universities. A three-year implementation has been agreed upon with the University of Iowa starting data collection activities in 2007.

[5]The log of a number can be computed in Excel using the "=LN" function (e.g., "=LN(3)" yields a log value of 1.09861). For the perversely curious, more details on logarithms can be found in Bruce Thompson (2006), <u>Foundations of Behavioral Statistics</u>, pages 403-407.

Figure 1
c:\p_point\arl_fig1.doc

Figure 2
c:\p_point\arl_fig2.doc

Q. *What was the goal of the present analyses?*
A. The goal of the analyses reported here was to develop some supplementary quantitative statistics that might be used by ARL libraries to help benchmark performance against additional quantitative statistics beyond the ARL Index. Conversely, the analyses were <u>not</u> conducted to offer alternative ARL membership criteria or to replace the ARL Index.

Q. *What quantitative data for ARL libraries were available for use in the present study?*
A. Only data collected by the ARL Statistics and Measurement Program were available for use in the analyses. Included are the five variables used in the current ARL Index.
Data from the years 2000 through 2004 were available at the time these analyses were performed. Data from multiple years were used to generate more stable results, by providing more data for selected analyses, and then to facilitate analyses that confirmed the stability of results across years. Results that are stable are inherently more useful in benchmarking efforts.
Data on several dozen variables have been collected during this 5 year time period. Additionally, data on an additional 8 variables, primarily involving digital content expenditures and library hours and number of staffed service points, have been collected only recently. Data on these last variables were widely available only for 2004. Table 1 lists all the variables available for the current study.

Q. *What was the primary statistical method used in the present analyses?*
A. The primary statistical method employed in these analyses is called factor analysis or principal component analysis. The purpose of factor analysis is to identify the groupings of variables that cluster together based on differentially larger relationships with each other[6]. These are the same statistical methods originally used in creating the ARL Index.

Q. *What were the initial analyses that were performed?*
A. The initial analyses were performed to begin to identify how many factors or components might be suitable for benchmarking, and which variables listed in Table 1 might be most suitable for these purposes. In selecting variables, some preference was afforded to using variables for which there was relatively little missing data. Some of these

---

[6]More detail on factor analysis is provided in Bruce Thompson (2004), <u>Exploratory and Confirmatory Factor Analysis</u>, which uses LibQUAL+® data for most of the book's heuristic examples.

initial analyses are presented in Appendix B.
Note that the sample sizes (i.e., $\underline{n}$'s) vary across different combinations of variables. A case was deleted if any data were missing for a given combination of variables. Thus, in general analyses involving more variables tended to have somewhat smaller sample sizes.

Table 1.
Variables Available for the Analyses

Abbreviation	Variable Label
Widely Available for 2000-2004	
vols	**'Vols In Library'**
volsadg	**'Vols Added (Gross)'**
volsadn	'Vols Added (Net)'
mono	'Monographs Purchased (Vols)'
serpur	'Curr Serials Purchased (Subs.)'
sernpur	'Curr Serials Not Purchased'
currser	**'total Current Serials'**
microf	'Microform Units'
govdocs	'Govt Documents'
compfil	'Computer Files'
mss	'Manuscripts and archives'
maps	'Carto-graphic Materials'
graphic	'Graphic Materials'
audio	'Audio Materials'
video	'Film / Video'
prfstf	'Prof Staff (FTE)'
nprfstf	'Support Staff (FTE)'
studast	'Stud Assistants (FTE)'
totstf	**'Total Staff W/O Students (FTE)'**
totstfx	'Total Staff W Students (FTE)'
expmono	'Monographs Expenditures'
expser	'Curr Serials Expenditures'
expoth	'Other Library Materials Expenditures'
expmisc	'Misc Materials Expenditures'
explm	'Tot Lib Materials Expenditures'
expbnd	'Contract Binding Expenditures'
salprf	'Prof Staff Salaries/Wages'
salnprf	'Supp Staff Salaries & Wages'
salstud	'Stud Asst Salaries & Wages'
totsal	'Total Salaries/Wages'
opexp	'Other Operating Expend'
totexp	**'Total Lib Expend'**
grppres	'Library Present to Groups'
presptcp	'Participants in Group Presentations'
reftrans	'Reference Transactions'
initcirc	'Initial Circ Transactions'
totcirc	'Total Circ Transactions'
illtot	'Loaned Total Items (ILL)'
ilbtot	'Borrowed Total Items (ILL)'
phdawd	'PhDs Awarded'
phdfld	'PhD Fields'
totstu	'Full-time (FTE) Total'
gradstu	'Grad Full-time (FTE)'
fac	'Faculty'

Widely Available Only for 2004
```
expcompf       'SU Computer Files Exp'
expeserl       'SU Elect Serials Exp.'
expbibul       'SU Lib Exp: Bibl Utilities, Networks'
expbibue       'SU Ext Exp: Bibl Utilities, Networks'
exphaso        'SU Exp: Computer Hardware Software'
expddill       'SU Exp: Doc Delivery/Interlib Loan Exp.'
svcpoint       'SU Staffed Service Points'
svchours       'SU Lib Service Hours'
```

_____

Note. The variables presented in **bold** are the 5 variables used in the current ARL Logarithmic Index.

Q. *What were the variables and factors that were isolated from these analyses?*

A. Table 2 presents the three components isolated in this analysis. The components involve Holdings, User Interactions, and Interlibrary Loan Activities. The components are reflected in 4, 3, and 2 measured variables respectively. The first component includes variables used in computing the ARL Index, although the Index employs log values of its variables, as explained previously.

Table 2.
Varimax-Rotated Principal Components for 9 Variables
Measured Across All 5 Years ($\underline{n}$ = 538)

Variable		Factor I	II	III
VOLS	'Vols In Library'	.92602	.22221	.13553
VOLSADG	'Vols Added (Gross)'	.91031	.23143	.08555
TOTSTF	'Total Staff W/O Students (FTE)'	.86895	.33393	.04720
CURRSER	'total Current Serials'	.85125	.21545	.16406
PRESPTCP	'Participants in Group Presentations'	.19965	.87510	.12552
GRPPRES	'Library Present to Groups'	.22521	.85142	.01234
REFTRANS	'Reference Transactions'	.32277	.61515	.17599
ILBTOT	'Borrowed Total Items (ILL)'	-.00867	.03965	.88865
ILLTOT	'Loaned Total Items (ILL)'	.27068	.18227	.76170

Note. Pattern/structure coefficients greater than $|.35|$ are underlined. The third and fourth eigenvalues (λ) were 1.14 and 0.62, respectively. The three components account for 78.6% of the observed variance in the 9 variables.

Q. *Must only variables already part of the ARL Statistics be used in these new library activity descriptions/statistics?*

A. No. First, the ARL Statistics and Assessment Committee could develop new variables that might flesh out these three dimensions (e.g., "unique/rare volumes held," "web-based reference transactions," "shared storage square feet," "collaboratively-held common holdings").
Second, new variables might be developed to define new dimensions not currently described by existing variables. For example, a Service Quality Improvement description might be developed, by adding variables such as "number of user focus

groups conducted," "FTE staff assigned to service quality assessment activities," "number of continuing education training sessions on service quality improvement attended by library staff").

Q. *Is the structure of these 9 variables sufficiently stable such that the scores on these 3 indices may be used across (a) institution types and (b) time?*

A. Yes, these components appear to be reasonably stable. Appendix C presents the component pattern/structure coefficients for different university types. Appendix D presents the component pattern/structure coefficients computed independently for each of the years 2000 through 2004.

    The coefficients also appear to be invariant to the use of different factor analytic computation methods. For example, Appendix E presents the pattern/structure coefficients computed using principal axis factor analysis, rather than principal components analysis.

Q. *How much do the scores on these three components overlap with scores on the existing ARL Index, and with other variables?*

A. Pearson product-moment correlation coefficients were computed between scores on the three components and scores on the ARL Index and other variables, and are reported in Appendix F.[7] When squared, these coefficients quantify the proportion of information (i.e., variability) that two scores have in common. Thus, r values greater than $|0.71|$ ($0.71^2 = 50\%$) indicate scores that have more than half their information in common.

    Scores on the first component, Holdings, and the ARL Index have 82.6% ($0.9090^2 = 0.826$) of their information in common. Of course, 82.6% $\neq$ 100%. The Holdings component also shares 74.7% ($0.8644^2 = 0.747$) of information with the variable, Total Library Expenditures. However, the User Interaction and the Interlibrary Loan indices have relatively little information in common with either the ARL Index or the other variables.

Q. *Is the three component structure stable when controlling for expenditure differences across libraries?*

A. The variables measuring or related to monetary expenditures were subjected to a principal components analysis in order to identify key expenditure variables. Two uncorrelated components emerged, as reported in Appendix G. The two key variables were Total Library Expenditures and Total Number of Full-time Equivalent Students.

    Two additional principal components analyses of the 9

---

[7]More information about correlation coefficients, and their properties and interpretations, can be found in Chapter 5 in Bruce Thompson (2006), <u>Foundations of Behavioral Statistics</u>.

variables were conducted using these two expenditure-related variables. First, all the variance in the 9 variables common to <u>Total Library Expenditures</u> was removed from the 9 variables, and then a principal components analysis of what was left in the 9 variables was conducted. Second, all the variance in the 9 variables common to <u>Total Number of Full-time Equivalent Students</u> was removed from the 9 variables, and then a principal components analysis of what was left in the 9 variables was conducted. The resulting components are presented in Appendix H. The results indicate that the Table 2 structure is stable even when controlling for key expenditure variables.

Q.   *How can these results be used by libraries for benchmarking purposes?*

A.   The three indices are **uncorrelated or independent** of each other. This means that scores on any combinations of the three indices may be used. Different libraries may reasonably focus on different combinations of indices.

For example, a given institution might elect to focus on their ARL Index rankings, but also look at standing on the User Interaction index. An interest in User Interaction dynamics is reasonable, given that our LibQUAL+® research shows that users care a lot about the service orientation and customer care focus of library staff.

The third statistic, Interlibrary Loan Activities, is a measure of the external connectedness of the library: how much a library contributes to and takes from the collection resources of the broader community. As reported in Appendix F, this connectedness is <u>not</u> driven by library wealth. For example, the common variance between Interlibrary Loan Activities scores and Total Library Expenditures is only 0.3% ($r^2 = 0.0565^2$).

One useful way to interpret these statistics invokes norms tables[8]. Appendix J presents some related normative tables. Tables such as these quantify what percentage of libraries fall below a given score. For example, in 2004, 75% of ARL libraries that year had a lower score on the Interlibrary Loan Expenditures index than the score of +0.233. The comparability of the normative tables across years for a given factor also supports the view that the indices may reasonably be used over time, and are not idiosyncratic in different years.

Q.   *How will the increasing movement toward the use of digital content affect the use of the ARL Index and the three supplementary indices described here?*

---

[8]The use of norms tables in the library context is explained in our <u>portal</u> journal article, Colleen Cook, Fred Heath, and Bruce Thompson (2002), "Score norms for improving library service quality: A LibQUAL+™ study" (vol. 2, pp. 13-26):
http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v002/2.1cook.html

A.   Unfortunately, the ARL Statistics and Measurement Program has
     only collected data related to digital content during the
     recent past. Thus, it was not possible to conduct a thorough
     statistical analysis as to whether these data will merely
     reflect the same patterns occurring within more traditional
     data. However, the correlation coefficients reported in
     Tables 3 and 4 offer some tentative insights into these
     issues.

Table 3.
Correlations of the New Indices with **Digital Spending**

Digital Spending	New Indices		
	Holdings	Interact	Loan_Act
EXPCOMPF	.2526	.0695	-.0846
	(97)	(97)	(97)
EXPESERL	.3163	.2545	.1520
	(105)	(105)	(105)
EXPBIBUL	.3792	.1147	.1001
	(100)	(100)	(100)
EXPBIBUE	.0660	.2991	.1307
	(50)	(50)	(50)
EXPHASO	.4439	.3377	.0342
	(105)	(105)	(105)
EXPDDILL	.3018	.0188	.1572
	(102)	(102)	(102)

Table 4.
Correlations of **Digital Spending** with Traditional Variables

Digital Spending	Traditional Variables			
	TOTEXP	TOTSTU	PHDAWD	INDEX
EXPCOMPF	.5561	.0960	.1990	.4108
	(102)	(102)	(101)	(102)
EXPESERL	.4151	.2559	.3016	.4088
	(110)	(110)	(109)	(110)
EXPBIBUL	.6721	.1294	.2698	.5091
	(105)	(105)	(104)	(105)
EXPBIBUE	.2335	.2493	.2672	.2221
	(52)	(52)	(51)	(52)
EXPHASO	.5865	.3103	.4807	.5757
	(110)	(110)	(109)	(110)
EXPDDILL	.2559	.3141	.5108	.3236
	(107)	(107)	(106)	(107)
SVCPOINT	.7339	.5490	.5670	.7393
	(111)	(111)	(110)	(111)
SVCHOURS	.2813	.1751	.2916	.2959
	(112)	(112)	(111)	(112)

First, the correlation coefficients reported in Table 3 suggest that the three indices proposed here overlap minimally with expenditures related to providing users access to digital content. This in turn suggests the potential for eventually developing a fourth statistic related to this service area.

Second, the results presented in Table 4 suggest some relationships between expenditures on digital content and aggregate expenditures. For example, there is 45.2% (r^2 = 0.6721^2) common information between Total Library Expenditures and Library Expenditures on Bibliographic Utilities and Networks. There is 34.4% (r^2 = 0.5865^2) common information between Total Library Expenditures and Library Expenditures on Computer Hardware and Software.

Of course, the "Google-ization" of the information world has brought exponentially increasing changes with research libraries, and even in the ways that library users think about libraries and librarians. A classic example of these impacts is this statement made by one of the users interviewed in grounding LibQUAL+® within the mindsets of users:[9]

> ...first of all, I would turn to the best search engines that are out there. That's not a person so much as an entity. In this sense, librarians are search engines [just] with a different interface.

Thus, the impacts of this evolution ought to be revisited on an on-going basis by the ARL Statistics and Measurement Program.

Q. *What form might a modified ARL Index statistic take?*
A. As University of Georgia University Librarian and Associate Provost William Potter noted in an e-mail communication with Martha Kyrillidou, the ARL Director of Statistics and Service Quality Programs, in December, 2005:

> Once you include one form of electronic publication in the "traditional" counts of volumes held and volumes added [in the ARL Index], then I do not see the logic behind not including other forms of electronic publication. For example, if a library scans a book itself and provides access to that book through the catalog, how is this different from the same book that a vendor has scanned and sold to the library? The really interesting thing in all this is that *if you have two libraries of similar nature and budget and one spends all its resources on print while the other is very aggressive in replacing print with*

---

[9]A comprehensive presentation of these qualitative studies is provided in the Ph.D. dissertation of Carol Colleen Cook (2001), "A mixed-methods approach to the identification and measurement of academic library service quality constructs: LibQUAL+™" (University Microfilms No. AAT3020024).

> *electronic, I expect that the library that focuses on print will have <u>higher ranks</u> in the ARL statistics even though the library users will have <u>less access</u>.* [emphasis added]

The use of a measure of <u>total expenditures</u> versus the use of some combination of (a) volume counts (historically part of the older statistics) and (b) expenditures on digital resources (only recently measured as part of the supplementary statistics) could (1) finesse the difficulty of distinguishing these two resources (2) while at the same time recognizing the changing face of the library in an increasingly digital world. Table 5 presents the pattern/structure coefficients for a component/factor involving two of the variables in the current ARL Index (total expenditures and total staff) and two expenditure variables (professional salaries and materials expenditures). Scores ($\underline{n}$ = 563) on the alternative factor from the correlation matrix (i.e., the rightmost factor) correlate $\underline{r}$ = +0.8997 ($\underline{r}^2$ = 0.8997^2 = 80.9%) with scores on the current ARL Index.

Table 5
An Expenditure-Focused Alternative Statistic

| | | Component Source | |
Variable		Covariance	Correlation
TOTEXP[a]	'Total Lib Expend'	1.000	.99236
TOTSTF[a]	'Total Staff w/o Students'	0.946	.96722
SALPRF[b]	'Prof Staff Salaries/Wages'	0.940	.94793
EXPLM[b]	'Total Lib Materials Expenditures'	0.944	.94408

[a]Already in the current ARL Index.
[b]<u>Not</u> in the current ARL Index.
<u>Note</u>. Factors or components can be computed from several sources, including a variance/covariance matrix, such that factors are sensitive to <u>both</u> relationships and data dispersion, or a correlation matrix, such that factors are based only on relationships among and the shapes of the data. Here the component structure was very similar across both computational methods, which means that the results were not an artifact of analytic choice.

Q. *Can <u>statistics</u> like those reported here make the decision about what variables should be built into quantitative Indices of library quality, or must the <u>people</u> within the ARL community instead make this decision?*

A. The responsibility for making the decision about what variables should be built into quantitative Indices of library quality can not atavistically be foisted onto statistics. Statistics can only help to inform these decisions by people. Instead, for two reasons, the <u>people</u> within the ARL community ultimately must make this decision.

First, analysis of the variables on which ARL has previously collected data is inherently limited, because these variables are neither (a) all the possible data about libraries nor (b) a random sample from the universe of all the possible choices of library datapoints. Thus, the decisions about what data were previously collected inescapably limit the generalizability of statistical results using the data. We simply could not use statistical analyses to build indices using data that were not even collected, for whatever reasons the absent data were not collected.

Second, statistical analyses of existing data tend to be backward-looking in perspective, rather than forward-looking. Unfortunately, given the rapid, transformational changes occurring with the world of information services, a forward-looking perspective is exactly what is required. Only people are well suited to anticipating what the library of tomorrow will look like.

In formulating a vision of the library of tomorrow, we will probably be well advised to remember the widely accepted wisdom that in predicting change we (a) tend to overestimate technological changes that will occur within a year and (b) massively underestimate the magnitude of changes that will occur over the course of the next 10 years. One solace is that our vision, once formulated, need not be fixed in form for all time; we can revisit our formulations as we continue to receive new information.