ARL Index and Other Validity Correlates of LibQUAL+™ Scores

Fred Heath, Colleen Cook, Martha Kyrillidou, and Bruce Thompson

Abstract: The present study was conducted to address five research questions bearing upon the psychometric validity of LibQUAL+™ scores using data provided by 20,416 participants. It was found that LibQUAL+™ subscale and total scores were highly correlated with satisfaction scores in two independent subsamples. As expected, respondents who reported never using the library systematically rated services lower than did other users. Also as expected, LibQUAL+™ mean scores—intended primarily to measure perceived service quality—were little correlated with institutional ARL Index scores.

The "New Measures" initiatives of the Association of Research Libraries (ARL) focus primarily on inputs such as library collection counts, but rather on outcome measures, such as assessments of expectations and perceptions of library service quality and user satisfaction. One "New Measures" initiative has been the LibQUAL+™ project.²

During the 2000–2001 academic year, termed the project "phase one" year in our ARL/Texas A&M University U.S. Department of Education Fund for the Improvement of Postsecondary Education (FIPSE) grant, the initial LibQUAL+™ protocol was refined down to twenty-five items. This process and the resulting product were described in detail in the two preceding articles in this issue.

LibQUAL+™ proposes to measure the perceived quality of library services in academic libraries. The first order of business is to assess the validity of the LibQUAL+™ scores to determine whether or not the instrument does what it purports to do. For

For some time library administrators, confronting growing fiscal constraints, have sought help in determining the most effective allocation of resources.


Reprinted with permission of The Johns Hopkins University Press.
some time library administrators, confronting growing fiscal constraints, have sought help in determining the most effective allocation of resources. In a psychometric context, validity deals with the issue of whether or not a measurement assesses the intended constructs related to these decisions. It has been suggested that “validity is the most important consideration in test evaluation.”4 If it can be shown that the test scores are valid, directors will be able to consider the merits of allocating resources among the four dimensions of library service (service affect, library as place, personal control, and information access) in order to satisfy user expectations.

Historically, test scores have been considered valid to the degree that it is known what the scores measure or predict.4 A more modern view is that validity is an evaluative judgment of “the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions” that are based on scores.5

The present study was conducted to address five questions regarding the validity of LibQUAL+™ scores. These questions were:

1. How well do LibQUAL+™ subscale (i.e., Service Affect, Library as Place, Personal Control, and Information Access) and total scores correlate with external validity scores (e.g., user ratings of service and satisfaction)?
2. Which of the 25 LibQUAL+™ item scores most differentiate the forty-three institutional affiliations of the 20,416 study participants?
3. Do mean ratings of perceived library service quality, as measured by LibQUAL+™ total T scores, differ with frequency of library use?
4. Do mean ratings of perceived library service quality, as measured by LibQUAL+™ total T scores, differ across user types (e.g., faculty members, graduate students)?
5. To what extent are institutional mean LibQUAL+™ subscale and total scores correlated with ARL Membership Criteria Index scores of the thirty-five participating libraries belonging to ARL?

In many ways this last question cuts to the core of the construct of effective library service quality. Wherever input metrics prevail as a dominating measure, the tendency is to concentrate resources on the algorithmic variables that are most influential in producing desired outcomes. Size of staff and expenditures for resources are input variables that exert tremendous influence in the ARL Index scores. A taut relationship between the Index and the LibQUAL+™ scores could suggest that Service Affect, Library as Place (a capital expenditure function that lies outside the reporting mechanism of the Index), and Personal Control remain subordinate to Information Access as areas of concern. Based on that premise, directors could opt to channel available new resources to that purpose rather than to eliminating deficits in the other three service dimensions.

Methods

Participants

As described in detail in the two preceding articles, the items on the web-administered LibQUAL+™ protocol were completed by 20,416 participants from forty-three campuses in the spring of 2001. Of these forty-three libraries, thirty-five were members of ARL.
Instrumentation

The participants rated their libraries on a 1-to-9 response format on the LibQUAL+™ items. This is the same response format used on the SERVQUAL protocol, on which LibQUAL+™ is based. Of course LibQUAL+™ incorporates some items unique to the library service context. These items were based on the multi-campus, multi-user-group interviews conducted by Cook and Heath.

Results

Question #1

The study’s first research question was, “How well do LibQUAL+™ subscale (i.e., Service Affect, Library as Place, Personal Control, and Information Access) and total scores correlate with external validity scores (e.g., user ratings of service and satisfaction)?” To address this question, LibQUAL+™ scale and total scores were converted to T scores with means of fifty and standard deviations of 10. The T scores are created by adding scores on items relevant to a given subscale (e.g., Service Affect), dividing by the number of items summed for a given score (e.g., 5, 9), and then standardizing to achieve the desired mean and standard deviation. These T scores were then correlated with three other sets of scores.

First, factor scores were computed from a factor analysis of the LibQUAL+™ data. Factor scores involve a mathematical weighting of the item response data more sophisticated than mere summation of item scores. The ultimate application of LibQUAL+™ scores involves standardized scores, such as T scores or percentile rank scores, but because the identification of the scales was partially based on factor analysis, it is hoped that the LibQUAL+™ T scores and factor scores would be highly related, thus showing the congruence of results across both score formats.

Second, participants responded to five library locus-of-control items evaluating whether they felt information acquisition resulted from personal skills or from chance dynamics. These five items are not part of LibQUAL+™ but were included in this one administration of the protocol only to assist us in better understanding what LibQUAL+™ does and does not measure (i.e., correlates of LibQUAL+™ scores). For our data the alpha reliability coefficient for these two very brief locus-of-control scales were .71 and .82, respectively, indicating that these scores were suitable for use in further analysis.

Third, the participants responded to three questions regarding their overall perceptions of service quality and satisfaction at their libraries:

1. “In general, I am satisfied with the way in which I am treated at the libraries”
   (1 = ‘strongly disagree’; 9 = ‘strongly agree’);
2. “In general, I am satisfied with library support for my learning, research and teaching needs”
   (1 = ‘strongly disagree’; 9 = ‘strongly agree’); and
3. “How would you rate the overall quality of the service provided by the library?”
   (1 = ‘extremely poor’; 9 = ‘extremely good’).
Library directors need to know that LibQUAL+™ performs well across a range of libraries.

When combined into a single scale, the reliability of these scores was .89.

We computed correlations among these variables separately for the 16,918 participants at ARL libraries and the 3,498 participants at non-ARL libraries. We did so to confirm our hypothesis that these results would replicate across library settings. Library directors need to know that LibQUAL+™ performs well across a range of libraries. Ideally, the measure may also be useful in smaller, non-ARL libraries. The resulting two correlation matrices are both presented in Table 1.

Question #2

The study's second research question was, "Which of the 25 LibQUAL+™ item scores most differentiate the forty-three institutional affiliations of the 20,416 study participants?" To address this research question we conducted a descriptive discriminant analysis.11

Descriptive discriminant analysis (DDA) identifies those items that most distinguish groups (i.e., here forty-three institutions) from each other. If no group differences can be isolated by this technique, then there simply are no group differences. In essence, such a finding would suggest that the forty-three libraries, an amalgam of ARL institutions from across the rankings of the ARL index, several research universities outside of ARL, and smaller universities with non-research missions were all one undifferentiated cohort. If every library was rated the same on all items, then LibQUAL+™ scores would not be useful in characterizing service quality or in identifying potential service quality improvements.

Table 2 presents the coefficients derived from this analysis. The standardized function coefficients presented in the table are analogous to regression beta weights, are used in DDA to derive the composite latent variables describing group differences (i.e., here the forty-three different libraries). The structure coefficients (r) presented in the tables are the correlations between the twenty-five measured variables and the discriminant function scores.

Descriptive discriminant analysis suggested that there were indeed perceived differences among the forty-three participating libraries. Together the three most noteworthy DDA functions (i.e., equations) that best described institutional differences accounted for about 57.5 percent of the explained variance. The canonical (i.e., multivariate) correlations associated with the three functions were .293, .269, and .235; larger correlations reflect the presence of greater differences in perceptions of the different academic libraries. The first three multivariate statistical significance tests were all statistically significant at α = .0001.

The first function was particularly useful in describing perceived differences between universities such as Brigham Young University, Cornell University, and the University of Illinois versus non-ARL libraries. As reported in Table 2, this function primarily involved six of the twenty-five items, and in particular items involving comprehensiveness of both print collections and journal runs. Here the collection strengths of these three ARL libraries make this distinction understandable.
Table 1
Product-moment Correlation Coefficients for LibQUAL+™ T and Factor Scores (v = 25) for (n = 16,918) ARL and (n = 3,498) non-ARL participants

<table>
<thead>
<tr>
<th>Score</th>
<th>TA_AFFEC</th>
<th>TA_PLACE</th>
<th>TA_PERCO</th>
<th>TA_INFOA</th>
<th>TA_LTOTL</th>
<th>LIBQ1</th>
<th>LIBQ2</th>
<th>LIBQ3</th>
<th>LIBQ4</th>
<th>LOCUS1</th>
<th>LOCUS2</th>
<th>SAT1</th>
<th>SAT2</th>
<th>SERVICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TA_AFFEC</td>
<td>1.0000</td>
<td>.48</td>
<td>.68</td>
<td>.60</td>
<td>.88</td>
<td>.91</td>
<td>.20</td>
<td>.26</td>
<td>.26</td>
<td>.37</td>
<td>-.33</td>
<td>.73</td>
<td>.60</td>
<td>.70</td>
</tr>
<tr>
<td>TA_PLACE</td>
<td>.4709</td>
<td>1.0000</td>
<td>.49</td>
<td>.53</td>
<td>.75</td>
<td>.21</td>
<td>.94</td>
<td>.15</td>
<td>.23</td>
<td>.29</td>
<td>-.12</td>
<td>.36</td>
<td>.44</td>
<td>.47</td>
</tr>
<tr>
<td>TA_PERCO</td>
<td>.6674</td>
<td>.4834</td>
<td>1.0000</td>
<td>.69</td>
<td>.85</td>
<td>.38</td>
<td>.21</td>
<td>.82</td>
<td>.37</td>
<td>.44</td>
<td>-.35</td>
<td>.54</td>
<td>.60</td>
<td>.66</td>
</tr>
<tr>
<td>TA_INFOA</td>
<td>.6128</td>
<td>.5238</td>
<td>.6934</td>
<td>1.0000</td>
<td>.82</td>
<td>.28</td>
<td>.26</td>
<td>.26</td>
<td>.88</td>
<td>.34</td>
<td>-.21</td>
<td>.44</td>
<td>.60</td>
<td>.62</td>
</tr>
<tr>
<td>TA_LTOTL</td>
<td>.8776</td>
<td>.7435</td>
<td>.8478</td>
<td>.8175</td>
<td>1.0000</td>
<td>.61</td>
<td>.48</td>
<td>.44</td>
<td>.47</td>
<td>.44</td>
<td>-.32</td>
<td>.65</td>
<td>.68</td>
<td>.75</td>
</tr>
<tr>
<td>LIBQ1</td>
<td>.9122</td>
<td>.2099</td>
<td>.3601</td>
<td>.3049</td>
<td>.6131</td>
<td>1.0000</td>
<td>.00</td>
<td>.02</td>
<td>.00</td>
<td>.24</td>
<td>-.27</td>
<td>.65</td>
<td>.41</td>
<td>.52</td>
</tr>
<tr>
<td>LIBQ2</td>
<td>.2024</td>
<td>.9422</td>
<td>.2159</td>
<td>.2556</td>
<td>.4826</td>
<td>.0007</td>
<td>1.0000</td>
<td>-.01</td>
<td>.03</td>
<td>.16</td>
<td>.00</td>
<td>.16</td>
<td>.23</td>
<td>.25</td>
</tr>
<tr>
<td>LIBQ3</td>
<td>.2509</td>
<td>.1628</td>
<td>.8287</td>
<td>.2847</td>
<td>.4416</td>
<td>-.0053</td>
<td>.024</td>
<td>1.0000</td>
<td>.00</td>
<td>.30</td>
<td>-.27</td>
<td>.26</td>
<td>.32</td>
<td>.34</td>
</tr>
<tr>
<td>LIBQ4</td>
<td>.2495</td>
<td>.1967</td>
<td>.3553</td>
<td>.8586</td>
<td>.4427</td>
<td>.0064</td>
<td>-.0160</td>
<td>.0021</td>
<td>1.0000</td>
<td>.19</td>
<td>-.08</td>
<td>.16</td>
<td>.40</td>
<td>.37</td>
</tr>
<tr>
<td>LOCUS1</td>
<td>.3319</td>
<td>.2950</td>
<td>.4326</td>
<td>.3326</td>
<td>.4178</td>
<td>.1915</td>
<td>.1815</td>
<td>.3087</td>
<td>.1769</td>
<td>1.0000</td>
<td>.00</td>
<td>.38</td>
<td>.40</td>
<td>.45</td>
</tr>
<tr>
<td>LOCUS2</td>
<td>-.2848</td>
<td>-.1034</td>
<td>-.3242</td>
<td>-.2021</td>
<td>-.2830</td>
<td>-.2212</td>
<td>-.0015</td>
<td>-.2567</td>
<td>-.1813</td>
<td>.0005</td>
<td>1.0000</td>
<td>-.29</td>
<td>.29</td>
<td>.32</td>
</tr>
<tr>
<td>SAT1</td>
<td>.7195</td>
<td>.3658</td>
<td>.5153</td>
<td>.4558</td>
<td>.6523</td>
<td>.6418</td>
<td>.1716</td>
<td>.2213</td>
<td>.1769</td>
<td>.3725</td>
<td>-.2700</td>
<td>1.0000</td>
<td>.64</td>
<td>.71</td>
</tr>
<tr>
<td>SAT2</td>
<td>.6065</td>
<td>.4258</td>
<td>.6041</td>
<td>.5829</td>
<td>.6733</td>
<td>.4237</td>
<td>.2277</td>
<td>.3330</td>
<td>.3561</td>
<td>.4272</td>
<td>-.2811</td>
<td>.6659</td>
<td>1.0000</td>
<td>.80</td>
</tr>
<tr>
<td>SERVICE</td>
<td>.6949</td>
<td>.4725</td>
<td>.6430</td>
<td>.6069</td>
<td>.7406</td>
<td>.5183</td>
<td>.2588</td>
<td>.3352</td>
<td>.3406</td>
<td>.4379</td>
<td>-.3004</td>
<td>.7329</td>
<td>.8027</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Note. To distinguish results for the two subsamples, coefficients for ARL participants are presented to 4 decimal places; coefficients for non-ARL participants are presented to 2 decimal places. Concurrent validity coefficients between LibQUAL+™ T scores and factor scores are italicized. Coefficients involving the LibQUAL+™ total scale T scores are presented in bold.
Table 1 Cont.

- "LIBQ1" = Service Affect factor scores. "LIBQ2" = Library as Place factor scores. "LIBQ3" = Personal Control factor scores. "LIBQ4" = Information Access factor scores.
- "LOCUS1" = Internal locus of control factor score; "LOCUS2" = Chance locus of control factor score.
- "SAT1" = "In general, I am satisfied with the way in which I am treated at the libraries" (1 = 'strongly disagree'; 9 = 'strongly agree'); "SAT2" = "In general, I am satisfied with library support for my learning, research and teaching needs" (1 = 'strongly disagree'; 9 = 'strongly agree');
- "SERVICE" = “How would you rate the overall quality of the service provided by the library?" (1 = 'extremely poor'; 9 = 'extremely good').

Table 2
Descriptive Discriminant Analysis Coefficients for Predicting Affiliation of (n = 20,416) Participants Using the 25 LibQUAL+™ Perceived Service Quality Items

<table>
<thead>
<tr>
<th>Item Core/Scale</th>
<th>Function I Func. r_s</th>
<th>Function II Func. r_s</th>
<th>Function III Func. r_s</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Service Affect</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Willingness to help users</td>
<td>.15 .18</td>
<td>-.16 .07</td>
<td>.08 .07</td>
</tr>
<tr>
<td>Giving users individual attention</td>
<td>-.04 .16</td>
<td>.08 .13</td>
<td>.02 .06</td>
</tr>
<tr>
<td>Employees who deal in a caring fashion</td>
<td>.10 .22</td>
<td>.23 .21</td>
<td>-.07 .02</td>
</tr>
<tr>
<td>Employees who are consistently courteous</td>
<td>.12 .23</td>
<td>.06 .19</td>
<td>-.10 .02</td>
</tr>
<tr>
<td>Employees have knowledge to answer</td>
<td>-.26 .13</td>
<td>-.24 .04</td>
<td>.12 .08</td>
</tr>
<tr>
<td>Employees who understand the needs</td>
<td>-.20 .16</td>
<td>-.16 .08</td>
<td>.00 .02</td>
</tr>
<tr>
<td>Readiness to respond to users' questions</td>
<td>.05 .19</td>
<td>-.09 .11</td>
<td>.16 .10</td>
</tr>
<tr>
<td>Employees who instill confidence</td>
<td>-.12 .16</td>
<td>.00 .15</td>
<td>-.05 .03</td>
</tr>
<tr>
<td>Dependability in handling users' needs</td>
<td>-.03 .25</td>
<td>-.04 .18</td>
<td>.07 .06</td>
</tr>
<tr>
<td><strong>Library as Place</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A haven for quiet and solitude</td>
<td>-.17 .26</td>
<td>-.26 .46</td>
<td>.43 .41</td>
</tr>
<tr>
<td>A meditative place</td>
<td>.18 .33</td>
<td>.42 .63</td>
<td>.24 .38</td>
</tr>
<tr>
<td>A contemplative environment</td>
<td>.05 .35</td>
<td>.21 .55</td>
<td>.00 .31</td>
</tr>
<tr>
<td>Space that facilitates quiet study</td>
<td>-.17 .23</td>
<td>.12 .49</td>
<td>.01 .27</td>
</tr>
<tr>
<td>A place for reflection and creativity</td>
<td>.16 .38</td>
<td>.47 .63</td>
<td>-.01 .27</td>
</tr>
</tbody>
</table>
The second function was particularly useful in describing perceived differences between universities on six items, as reported in Table 2, including all five LibQUAL+™ items measuring the Library as Place (e.g., “A place for reflection and creativity”). While further information relating to the specific institutions is required, it appears that there may be noteworthy distinctions between those libraries with aged or inadequate physical plants and those that have recently undergone successful new construction/renovation. The third function involved perceived differences on three items, including convenient business hours. These functions suggest perceived differences in both the qualities of physical spaces and expectations for access, reflecting differences at the local level.

**Question #3**

The study’s third research question was, “Do mean ratings of perceived library service quality, as measured by LibQUAL+™ total T scores, differ with frequency of library use?” That is, has the survey discovered anything about the perceptions of frequent library users that may be different from perceptions of infrequent users or nonusers? Is it possible that frequent users are more demanding, or that the responses of those in the latter category may reflect users who are alienated or in some other way underserved? We compared scores across use of both the physical library and electronic use.

Table 3 presents the LibQUAL+™ total means by categories of use frequency for the 16,918 participants from ARL campuses. Results for the non-ARL participants were comparable, and so are not presented here in the interest of saving space. The table presents both ANOVA F tests and 95 percent confidence intervals about each mean. Also presented are the η² effect sizes that are analogous to r².
Table 3

ANOVA of LibQUAL+™ Total T Scores by Frequency of Library Use

<table>
<thead>
<tr>
<th>Group</th>
<th>Count</th>
<th>Mean</th>
<th>(SD)</th>
<th>95% CI for Mean</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Library Use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>daily</td>
<td>2582</td>
<td>50.026</td>
<td>(10.6133)</td>
<td>49.6167 to 50.4358</td>
<td>0.15%</td>
</tr>
<tr>
<td>weekly</td>
<td>7346</td>
<td>49.8425</td>
<td>(10.1402)</td>
<td>49.6106 to 50.0744</td>
<td></td>
</tr>
<tr>
<td>monthly</td>
<td>4772</td>
<td>50.2789</td>
<td>(9.4972)</td>
<td>50.0093 to 50.5484</td>
<td></td>
</tr>
<tr>
<td>quarter</td>
<td>2031</td>
<td>50.1815</td>
<td>(9.5648)</td>
<td>49.7653 to 50.5978</td>
<td></td>
</tr>
<tr>
<td>never</td>
<td>187</td>
<td>46.7372</td>
<td>(12.0638)</td>
<td>44.9969 to 48.4776</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>16918</td>
<td>50.0000</td>
<td>(10.0000)</td>
<td>49.8493 to 50.1507</td>
<td></td>
</tr>
</tbody>
</table>

$F_{calc} = 6.54; df = 4/16,913; p_{calc} = .000003$

Electronic Use Frequency

<table>
<thead>
<tr>
<th>Group</th>
<th>Count</th>
<th>Mean</th>
<th>(SD)</th>
<th>95% CI for Mean</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>daily</td>
<td>3468</td>
<td>49.0486</td>
<td>(10.6043)</td>
<td>48.6955 to 49.4016</td>
<td>0.29%</td>
</tr>
<tr>
<td>weekly</td>
<td>6836</td>
<td>50.1117</td>
<td>(9.9144)</td>
<td>49.8767 to 50.3468</td>
<td></td>
</tr>
<tr>
<td>monthly</td>
<td>3407</td>
<td>50.6584</td>
<td>(9.7518)</td>
<td>50.3308 to 50.9859</td>
<td></td>
</tr>
<tr>
<td>quarter</td>
<td>1901</td>
<td>50.2566</td>
<td>(9.5376)</td>
<td>49.8276 to 50.6857</td>
<td></td>
</tr>
<tr>
<td>never</td>
<td>1306</td>
<td>49.8506</td>
<td>(9.9186)</td>
<td>49.3121 to 50.3890</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>16918</td>
<td>50.0000</td>
<td>(10.0000)</td>
<td>49.8493 to 50.1507</td>
<td></td>
</tr>
</tbody>
</table>

$F_{calc} = 12.17; df = 4/16,913; p_{calc} = 7.02 \times 10^{-10}$

Question #4

The study’s fourth research question was, “Do mean ratings of perceived library service quality, as measured by LibQUAL+™ total T scores, differ across user types (e.g., faculty members, graduate students)?” This concern is an analog of Question #3. Is it possible that different user groups view us differently, based upon needs or priorities that may vary among the primary groups? Table 4 presents the ANOVA $F$ tests, 95 percent confidence intervals about each mean, and the $\eta^2$ effect sizes from the analysis involving the 16,918 participants from ARL campuses.

Question #5

As we have suggested, the fifth research question, whose resolution is most anticipated by library managers and other campus administrators, was: “To what extent are institutional mean LibQUAL+™ subscale and total scores correlated with ARL Membership Criteria Index scores of the 35 participating libraries belonging to ARL?” The ARL Membership Criteria Index score is a summary measure of the relative size of investments in or resources of university libraries; the sole purpose of the Index is to determine institutional membership eligibility to be admitted to the ARL.\textsuperscript{13}
Table 4
ANOVA of LibQUAL+™ T Scores by User Types

<table>
<thead>
<tr>
<th>Group</th>
<th>Count</th>
<th>Mean</th>
<th>(SD)</th>
<th>95% CI for Mean</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Service Affect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Library</td>
<td>743</td>
<td>51.1794</td>
<td>(8.9751)</td>
<td>50.5330 to 51.8258</td>
<td>0.74%</td>
</tr>
<tr>
<td>Other</td>
<td>1035</td>
<td>51.0456</td>
<td>(9.6009)</td>
<td>50.4600 to 51.6312</td>
<td></td>
</tr>
<tr>
<td>Faculty</td>
<td>4420</td>
<td>50.9012</td>
<td>(10.0318)</td>
<td>50.6054 to 51.1970</td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>5566</td>
<td>49.9878</td>
<td>(9.9865)</td>
<td>49.7254 to 50.2502</td>
<td></td>
</tr>
<tr>
<td>Undergr</td>
<td>5154</td>
<td>48.8603</td>
<td>(10.0884)</td>
<td>48.5849 to 49.1358</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>16918</td>
<td>50.0000</td>
<td>(10.0000)</td>
<td>49.8493 to 50.1507</td>
<td></td>
</tr>
<tr>
<td>F₁₅₂₇ = 31.35; df = 4/16,913; p₁₅₂₇ = 4.71 E-26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Library as Place</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undergr</td>
<td>5154</td>
<td>52.2390</td>
<td>(9.4249)</td>
<td>52.0816 to 52.5964</td>
<td>3.50%</td>
</tr>
<tr>
<td>Other</td>
<td>1035</td>
<td>52.1844</td>
<td>(8.9565)</td>
<td>51.6381 to 52.7307</td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>5566</td>
<td>49.3767</td>
<td>(9.9174)</td>
<td>49.1161 to 49.6372</td>
<td></td>
</tr>
<tr>
<td>Faculty</td>
<td>4420</td>
<td>48.0932</td>
<td>(10.2626)</td>
<td>47.7905 to 48.3958</td>
<td></td>
</tr>
<tr>
<td>Library</td>
<td>743</td>
<td>46.7453</td>
<td>(10.3203)</td>
<td>46.0020 to 47.4886</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>16918</td>
<td>50.0000</td>
<td>(10.0000)</td>
<td>49.8493 to 50.1507</td>
<td></td>
</tr>
<tr>
<td>F₁₅₂₇ = 153.44; df = 4/16,913; p₁₅₂₇ = 3.62 E-129</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Personal Control</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1035</td>
<td>50.5828</td>
<td>(9.5882)</td>
<td>49.9980 to 51.1676</td>
<td>0.20%</td>
</tr>
<tr>
<td>Graduate</td>
<td>5566</td>
<td>50.4993</td>
<td>(9.9960)</td>
<td>50.2366 to 50.7619</td>
<td></td>
</tr>
<tr>
<td>Undergr</td>
<td>5154</td>
<td>49.9164</td>
<td>(10.0694)</td>
<td>49.6414 to 50.1913</td>
<td></td>
</tr>
<tr>
<td>Faculty</td>
<td>4420</td>
<td>49.4435</td>
<td>(10.2030)</td>
<td>49.1426 to 49.7434</td>
<td></td>
</tr>
<tr>
<td>Library</td>
<td>743</td>
<td>49.3388</td>
<td>(8.5794)</td>
<td>48.7209 to 49.9567</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>16918</td>
<td>50.0000</td>
<td>(10.0000)</td>
<td>49.8493 to 50.1507</td>
<td></td>
</tr>
<tr>
<td>F₁₅₂₇ = 8.69; df = 4/16,913; p₁₅₂₇ = .00000005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Information Access</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Library</td>
<td>743</td>
<td>51.9045</td>
<td>(8.4042)</td>
<td>51.2992 to 52.5098</td>
<td>1.87%</td>
</tr>
<tr>
<td>Other</td>
<td>1035</td>
<td>51.8612</td>
<td>(9.2535)</td>
<td>51.2968 to 52.4256</td>
<td></td>
</tr>
<tr>
<td>Undergr</td>
<td>5154</td>
<td>51.1751</td>
<td>(9.0457)</td>
<td>50.9281 to 51.4221</td>
<td></td>
</tr>
<tr>
<td>Faculty</td>
<td>4420</td>
<td>47.9684</td>
<td>(10.7821)</td>
<td>47.6504 to 48.2863</td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>5566</td>
<td>49.9249</td>
<td>(10.2416)</td>
<td>49.6558 to 50.1940</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>16918</td>
<td>50.0000</td>
<td>(10.0000)</td>
<td>49.8493 to 50.1507</td>
<td></td>
</tr>
<tr>
<td>F₁₅₂₇ = 80.67; df = 4/16,913; p₁₅₂₇ = 6.18 E-68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4 Cont.

<table>
<thead>
<tr>
<th>Group</th>
<th>Count</th>
<th>Mean</th>
<th>(SD)</th>
<th>95% CI for Mean</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LibQUAL+™ Total Score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1035</td>
<td>51.6527</td>
<td>(9.6019)</td>
<td>51.0671 to 52.2384</td>
<td>0.39%</td>
</tr>
<tr>
<td>Undergr</td>
<td>5154</td>
<td>50.4450</td>
<td>(9.8675)</td>
<td>50.1756 to 50.7145</td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>5566</td>
<td>49.9274</td>
<td>(10.0258)</td>
<td>49.6639 to 50.1908</td>
<td></td>
</tr>
<tr>
<td>Library</td>
<td>743</td>
<td>49.7397</td>
<td>(9.0998)</td>
<td>49.0843 to 50.3951</td>
<td></td>
</tr>
<tr>
<td>Faculty</td>
<td>4420</td>
<td>49.2293</td>
<td>(10.2829)</td>
<td>48.9261 to 49.5325</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>16918</td>
<td>50.0000</td>
<td>(10.0000)</td>
<td>49.8493 to 50.1507</td>
<td></td>
</tr>
</tbody>
</table>

F calc = 16.44; df = 4/16,913; p calc = 1.88 E-13

While the Index score is not intended to measure a library’s services or success in meeting users’ needs, the relationship between input metrics and outcomes has long been intriguing. A groundbreaking issue of Library Trends first addressed this question. At that point the question of the relationship between expenditures and quality was joined, and the on-going discussion has had far-reaching implications, including in some respects the development of LibQUAL+™ and other New Measures initiatives. On the other hand, we anticipated a near-zero correlation of institutional LibQUAL+™ means with ARL Membership Criteria Index scores. The scores of the new instrument measure perceptions of service quality across the dimensions of Service Affect and Personal Control while taking into account the continuing importance of Library as Place. The additional element of Information Access (the "+" in LibQUAL+™) remains important but is not the dominating element that it is in the Index. Even libraries with very limited resources may provide dedicated service to users, and users probably adjust their service expectations according to what they deem reasonable in a given service setting. Table 5 presents the bivariate correlations of ARL Membership Criteria Index scores with both LibQUAL+™ gap scores and perception scores for both subscales and the total protocol.

The four LibQUAL+™ subscale scores were also used to predict ARL Membership Criteria Index scores of the thirty-five institutions in a multiple regression analysis. The R² from this analysis was .068. This is a low-to-moderate effect size (Thompson, in press) indicating that knowledge of the four LibQUAL+™ subscale means explains only 6.8 percent of the variability in Index scores of these thirty-five schools. This result reflects the near-zero correlation of the LibQUAL+™ and Index scores that we anticipated.

The beta weights and structure coefficients from this analysis are presented in table 6. As reported in table 6, the LibQUAL+™ subscale score most associated with the ARL Membership Criteria Index scores was Access to Collections (r = .564; β = .323). Thus, as might be expected, to the minimal extent that the four LibQUAL+™ subscale means somewhat predict differences in Index scores, most of this overlap involved the plus ("+") in LibQUAL+™: collections.
Table 5
Correlations of LibQUAL+™ Gap and Perceived Score with ARL Membership Criteria Index Scores (n = 35 ARL Institutions)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Gap Score Adequacy</th>
<th>Gap Score Superiority</th>
<th>LibQUAL+™ Score Affect</th>
<th>LibQUAL+™ Score Place</th>
<th>LibQUAL+™ Score Control</th>
<th>LibQUAL+™ Score Access</th>
<th>LibQUAL+™ Score Total</th>
<th>ARL Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adequacy</td>
<td>1.000</td>
<td>.920</td>
<td>.725</td>
<td>.680</td>
<td>.686</td>
<td>.637</td>
<td>.800</td>
<td>.208</td>
</tr>
<tr>
<td>Superiority</td>
<td>1.000</td>
<td>.832</td>
<td>.772</td>
<td>.756</td>
<td>.638</td>
<td>.888</td>
<td>.095</td>
<td></td>
</tr>
<tr>
<td>Perceived Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affect</td>
<td>1.000</td>
<td>.588</td>
<td>.782</td>
<td>.655</td>
<td>.897</td>
<td>.081</td>
<td>.043</td>
<td></td>
</tr>
<tr>
<td>Place</td>
<td></td>
<td></td>
<td>1.000</td>
<td>.606</td>
<td>.552</td>
<td>.822</td>
<td>.004</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>.739</td>
<td>.890</td>
<td>.004</td>
<td></td>
</tr>
<tr>
<td>Access</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>.820</td>
<td>.147</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>.048</td>
<td></td>
</tr>
</tbody>
</table>

Note. "Adequacy" = "Perceived" service quality - "Minimally acceptable" service quality. "Superiority" = "Perceived" service quality - "Desired" service quality. All coefficients, except those involving the ARL Index, are statistically significant (p < .001).

Table 6
LibQUAL+™ Gap and Perceived Score ARL Membership Criteria Index Scores Predicted by the Four LibQUAL+™ Perceived Subscale Scores (n = 35 ARL Institutions)

<table>
<thead>
<tr>
<th>LibQUAL+™ Scale</th>
<th>rs</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access to Collections</td>
<td>.564</td>
<td>.323</td>
</tr>
<tr>
<td>Personal Control</td>
<td>.015</td>
<td>-.281</td>
</tr>
<tr>
<td>Service Affect</td>
<td>.311</td>
<td>.182</td>
</tr>
<tr>
<td>Library as Place</td>
<td>-.165</td>
<td>-.158</td>
</tr>
</tbody>
</table>

Note: R = .2605; df = 4/30; p = .7302.
Discussion

The study participants were primarily library users, as reported in table 3. For example, 98.9 percent of the 16,918 participants at ARL institutions used the physical library at least quarterly. Similarly, 92.3 percent of the participants at ARL libraries used electronic resources at least quarterly. Fortunately, the participants were usually able to base their ratings on actual use experiences. These results suggest that since most members of university communities do use libraries, elaborate survey designs intended to reach non-users may be unnecessary. Overall, the answers to our five research questions seem to support a conclusion that LibQUAL+™ scores are reasonably valid.

The Table 1 validity coefficients are encouraging in at least three respects. It is especially striking that results replicated so closely across the independent ARL and non-ARL samples. This suggests the important conclusion that LibQUAL+™ scores may be valid in reasonably diverse library settings. Thus, a single instrument might be used to measure service quality across different types of post-secondary libraries.

First, the LibQUAL+™ T scores based on standardizing sums of item responses were highly correlated (e.g., .9122, .9422, .91, .94) with the factor scores on which the subscale definitions were based.¹⁸ This result suggests that the scores are not an artifact of scoring rubric, and that they are valid.

Second, the LibQUAL+™ T scores were positively correlated with Internal locus of control (i.e., belief that users themselves determine information seeking success) and negatively correlated with Chance locus of control (i.e., the belief that success in information seeking is a random function). Locus of control constructs have been elaborated in social learning theory.¹⁹ According to social learning theory, persons who believe that they control their own destinies, i.e., internals, behave in predictable ways in comparison with their external counterparts, i.e., persons who believe that chance or powerful others determine outcomes in their lives.

Not surprisingly, the LibQUAL+™ subscale most correlated (.4326 and .44 in the two samples, as reported in table 1) with the Internal locus of control scores was Personal Control. This would be hoped for from a score validity point of view, because among the four subscales this one on its face seems most relevant to having an internal locus of control. The results seem to validate a conclusion that library administrators would be wise to focus efforts upon improving those variables that enhance library users' feelings of personal control. Bibliographic instruction, information services, signage and other way-finding mechanisms and an improved web presence may be among such variables.

Third, all three satisfaction items presented at the end of the survey were highly correlated (e.g., .7406 and .75, respectively, for service scores) with LibQUAL+™ scores. Clearly there is a strong relationship between high ratings for overall quality of library service and user satisfaction. While this relationship is patently intuitive, it is important
to recall that all four dimensions of library service quality are simultaneously at work in influencing this relationship.

As further evidence of validity, perceptions of the Library as Place were least related (e.g., .4725 and .47, respectively) to ratings of satisfaction. Users apparently think about library as place in conceptualizing service quality (cf. Cook & Heath, 2001), but service satisfaction itself apparently is more related to ability to locate needed information.\(^{20}\)

The Table 3 results suggest that participants who “never” use the physical library were most negative in their mean perceptions ratings (46.7; SD = 12.1). The possible reasons for non-use, such as perceived resource inadequacy, prior service failure, or an impulsive reaction to a library-sponsored survey, are not addressed here. Intuitively, this finding is not surprising, and in its own way supports score validity. However, there are only minute differences in perceptions about electronic use. The reasons for never visiting the library virtually could be related more to issues of connectivity, personal preference or technical competence and less to issues of resource adequacy. For those users, physical access may suffice.

The Table 4 results suggest that in general there are not systematic differences in perceptions across user types. However, library staff did tend to rate Library as Place somewhat lower than did other participants. These results suggest that users may have similar perceptions of library service quality regardless of their vantage points, even though there are some differences in the service needs of these groups.\(^{21}\) In some ways, this finding may validate qualitative findings of the relationships between faculty and their students. The attitudes of faculty regarding library service quality are grounded in their own experiences with library use over their academic careers, and those expectations may be conveyed in their mentoring relationships with graduate students or in classroom discourse with undergraduates. From these relationships may emerge predispositions regarding library service quality, a concept that merits further study.

The Table 5 results show that means on perceived service adequacy are highly positively correlated (e.g., .725, .680 with “gap” means on service adequacy (i.e., perceived service adequacy minus minimally-acceptable service adequacy); thus institutions with favorable means on perceived scores also tended to have larger gaps involving perceptions exceeding minimum expectations. Institutional means of perceived service adequacy were also highly positively correlated (e.g., .832, .772, .756) with “gap” means on service superiority (i.e., perceived service minus desired service adequacy); thus, institutions with favorable means on perceived scores also tended to have smaller gaps involving perceptions approaching or exceeding desired service quality. Because perceptions themselves as well as gap scores are alternative ways of conceptualizing user views, all these results would be expected. However, the LibQUAL\(^{+}\) service adequacy gap means had a higher correlation (.208) with the ARL Membership Criteria Index scores of the thirty-five ARL institutions than did either the LibQUAL\(^{+}\) superiority gap means (.095) or LibQUAL\(^{+}\) total perception scores (.048).

In the aggregate, the LibQUAL\(^{+}\) means were not appreciably correlated with ARL Membership Criteria Index scores concerning the bivariate results reported in Table 5. The biggest correlation involved Information Access, as would be expected, because that subscale involves an element of collections, but even this correlation was small ($r^2 = .147^2 = 2.2\%$).
The same pattern occurred in the multiple regression analyses invoking the LibQUAL+™ means as a system to predict Index scores. The largest beta weight (.564) and structure coefficient (.323) occurred for Information Access. However, this occurred in the context of a multiple $R^2$ using all four subscale means that was only 6.8%, and the “adjusted” $R^2$ was negative.22 The first result indicates that the Index and the LibQUAL+™ subscale scores had low-to-moderate common variance, while the second result indicates that even this modest predictability might have been due to sampling error. During the 2001–2002 academic year LibQUAL+™ may be completed at 200 campuses, and at that juncture with a larger and even more diverse sample of institutions we will know a lot more about how much service quality scores and collection indices have in common with each other.

The results described above are important evidence of the validity of LibQUAL+™ scores. Effective organizational assessment requires multiple listening methods. LibQUAL+™ has been developed to supplement important input-based metrics, such as the ARL Index, but not to replace them or to provide the only listening method. An important assumption of LibQUAL+™ is that users behave rationally and adjust their expectations according to the mission of the university and to its particular local context.

It would be unexpected to find that only large, well-funded research universities provide good services across the dimensions that define the constructs in LibQUAL+™. Libraries at smaller universities with disparate missions can also provide dedicated service and give empathetic attention to users’ needs. Indeed, we have consistently found across diverse campuses that users have high expectations of service, and that within the local context their expectations are often reasonably well met. The usefulness of LibQUAL+™ is that in each of these contexts, regardless of the local mission, strengths and deficits can be identified, and strategies can be developed for improvement.

LibQUAL+™ provides the library administrator with an opportunity to meet the requirements for accountability. In the university setting, providing information resources for the academic community is a public good that enjoys strong consensus. Translated into the budgetary context, there is unrelenting pressure for additional resources in order to fulfill this purpose. If LibQUAL+™ can help library administrators demonstrate, and campus resource allocators understand, that additional resources effectively address all the issues pervading the construct of library service quality, then the task of advancing the library mission may become more sustainable. Across all sectors and all organizations, total market survey protocols provide insight into user or customer perceptions of return on investments. We believe that LibQUAL+™ is proving to be one important tool for better listening to users.

Fred Heath is Dean of Library Services and holder of the Sterling Evans Chair at Texas A&M University; he may be contacted via e-mail at: fheath@tamu.edu.

Colleen Cook is Professor of Library Science and Executive Associate Dean of the Sterling C. Evans Library at Texas A&M University; she may be contacted via e-mail at: ccook@tamu.edu.
Martha Kyrillidou is a Senior Program Officer, ARL Statistics and Measurement Program, Association of Research Libraries (ARL); she may be contacted via e-mail at: martha@arl.org.

Bruce Thompson is Professor of Educational Psychology and Distinguished Research Scholar, Texas A&M University, and Adjunct Professor of Community Medicine, Baylor College of Medicine; he may be contacted via Internet URL: <http://www.coe.tamu.edu/~btthompson>.

Notes
1. This research was supported in part by a grant from the U.S. Department of Education Fund for the Improvement of Postsecondary Education (FIPSE).
9. B. Thompson, Cook, and R. L. Thompson, “Reliability and Structure of LibQUAL+™ Scores” [THIS ISSUE].
16. B. Thompson, "'Statistical, 'Practical,' and 'Clinical.'"
18. Cook and others, "The ARL 'LibQUAL+™' Study—A Preliminary Report."
20. Cook and Heath, "A 'LibQUAL+™' Qualitative Interview Study."
21. Ibid.
22. B. Thompson, "'Statistical, 'Practical,' and 'Clinical.'"