



# Reliability and Structure of LibQUAL+™ Scores: Measuring Perceived Library Service Quality

---

Bruce Thompson, Colleen Cook and Russel L.  
Thompson

**abstract:** LibQUAL+™ is a tool designed to measure user perceptions of service quality in academic libraries. This web-based protocol was recently completed by 20,416 participants representing forty-three universities. For the purposes of the present report these data were evaluated to address two research questions: (a) how reliable were the LibQUAL+™ scores?; and (b) what were the dimensions underlying user perceptions of library service quality? The second question focuses upon identifying the most appropriate scales to report LibQUAL+™ scores to librarians using the protocol to improve service quality.

**A**cademic libraries today confront escalating pressure to demonstrate impact. In the contemporary campus milieu, "every unit . . . is valued in proportion to its contribution to the quality success of the campus."<sup>1</sup> But as Cullen recently noted,

Academic libraries are currently facing their greatest challenge since the explosion in tertiary education and academic publishing which began after World War II. The global digital revolution is affecting both the traditional forms of the creation, organization and dissemination of knowledge, and the world of tertiary education itself. The alliance of business and universities to create a new paradigm of tertiary education, and the emergence of the virtual university, supported by the virtual library, calls into question many of our basic assumptions about the role of the academic library, and the security of its future. Retaining and growing their customer base, and focusing more energy on meeting their customers' expectations is the only way for academic libraries to survive in this volatile environment.<sup>2</sup>

*portal: Libraries and the Academy*, Vol. 2, No. 1 (2002), pp. 3–12. Copyright © 2002 by The Johns Hopkins University Press, Baltimore, MD 21218.

Reprinted with permission of The Johns Hopkins University Press.

In this environment, "A measure of library quality based solely on collections has become obsolete."<sup>3</sup>

These various considerations have prompted the Association of Research Libraries (ARL) to sponsor a number of "New Measures" initiatives. The New Measures efforts represent a collective determination on the part of the ARL membership to augment the collection-count and fiscal *input* measures that comprise the ARL Index and ARL Statistics, which are to date the most consistently collected statistics for research libraries, with *outcome* measures such as assessments of service quality and satisfaction.

One New Measures initiative has been the "LibQUAL+™" project.<sup>4</sup> LibQUAL+™ grew out of initial efforts to apply the widely used "SERVQUAL" protocol in the library service context. The SERVQUAL instrument, which includes twenty-two items ostensibly measuring perceptions of *tangibles*, *reliability*, *responsiveness*, *assurance* and *empathy*, has been fairly widely used in business settings.<sup>5</sup> Within this model, "only customers judge quality; all other judgments are essentially irrelevant."<sup>6</sup>

SERVQUAL has also been applied with some frequency in library settings. However, these applications have consistently failed to recover the expected five score dimensions (e.g., *tangibles*, *reliability*) found in business applications.<sup>7</sup> Furthermore, given the context of its development in the business sector, SERVQUAL may not measure some service dimensions unique to the library setting, particularly electronic services, that may indeed be very important to users in these settings.

The 1999–2000 academic year constituted phase zero of the LibQUAL+™ project. This phase had two primary components. First, a qualitative study was conducted to regroup the LibQUAL+™ protocol in the university library setting.<sup>8</sup> Based on content analysis of user interviews conducted at nine university campuses, it was clear that additional items were needed to measure additional dimensions (e.g., the library as place, ubiquity of access, facilitation of user personal control) of service quality in higher education settings.

Second, a quantitative study was undertaken to explore how well an expanded measure performed in evaluating user perceptions of library service quality.<sup>9</sup> In the phase zero 1999–2000 academic year the web-based LibQUAL+™ tool consisting of forty-one (twenty-two SERVQUAL plus nineteen trial) items was completed by 4,407 participants from thirteen ARL institutions.

Although these studies yielded very promising results, additional item refinement and investigation was deemed appropriate during the LibQUAL+™ phase one 2000–2001 academic year. A primary purpose of this phase was to investigate the psychometric integrity of scores from the revised measure. Toward this end the revised web-based protocol was completed by 20,416 participants representing forty-three universities. Specifically, the present study was conducted to evaluate two research questions:

1. How reliable were the LibQUAL+™ scores generated in the present study?
2. What were the dimensions underlying user perceptions of library service quality?

The second question focuses upon identifying the most appropriate scales to report LibQUAL+™ scores to librarians using the protocol to improve service quality.

## Method

### Sample

The directors of the libraries at the forty-three universities solicited participation in the survey from random samples of potential library users on each campus. Table 1 presents the breakdown of the completed surveys obtained at each institution.

---

**Table 1**  
LibQUAL+™ Phase 1 Participating Schools

---

Category / University	<i>n</i>	%
<i>ARL Libraries (35)</i>		
University of Arizona	302	1.5
Arizona State University	252	1.2
Brigham Young University	789	3.9
Cornell University	634	3.1
Emory University	341	1.7
University of Guelph	794	3.9
University of Houston	304	1.5
University of Illinois	669	3.3
Indiana University	228	1.1
University of Iowa	426	2.1
Iowa State University	717	3.5
University of Kentucky	801	3.9
University of Maryland	236	1.2
McGill University	663	3.2
Michigan State University	320	1.6
University of Minnesota	443	2.2
University of Missouri	426	2.1
University of Nebraska	295	1.4
University of New Mexico	198	1.0
Galter Health Science	476	2.3
Ohio University	590	2.9
Oklahoma State University	208	1.0
University of Oregon	751	3.7
University of Pittsburgh	399	2.0
Southern Illinois University	287	1.4
University of Texas	697	3.4
Texas Tech University	442	2.2

Table 1 Cont.

<b>Category / University</b>	<i>n</i>	%
University of Colorado	230	1.1
University of Utah	397	1.9
Virginia Tech University	877	4.3
University of Washington	274	1.3
Washington State University	382	1.9
University of Waterloo	967	4.7
Wayne State University	270	1.3
Texas A&M University	833	4.1
<b>Subtotal</b>		<b>16918</b>
<i>non-ARL Universities (8)</i>		
American University	548	2.7
University of Arkansas	471	2.3
Baylor University	475	2.3
Clemson University	266	1.3
Miami University	1000	4.9
University of Mississippi	193	.9
Oregon State University	329	1.6
Washburn University	216	1.1
<b>Subtotal</b>		<b>3498</b>

The final sample included data from 6,084 undergraduate students (29.8 percent), 6,482 graduate students (31.7 percent), 5,599 faculty (27.4 percent), and 2,251 others, including library staff and other staff personnel (11.0 percent). The sample included 10,647 females (51.3 percent) and 9,949 males (48.7 percent).

### **Instrumentation**

The LibQUAL+™ survey is administered on the Web. Examples of the visual presentation of items are presented in our prior work.<sup>10</sup> Participants respond to items using a 1-to-9 Likert scale, which is the same response format used on the SERVQUAL protocol. Users click on selected response choices in what is called a radio button response format; clicking inside a response circle causes that circle to be filled and the response to be recorded accordingly.

### **Results**

#### **Score Reliability**

Score reliability is essential to any attempt to use data for service improvement. For example, if one is attempting to diet, and on a given morning stands on a bathroom scale and

finds upon repeated measurements that reported weights fluctuate wildly by 20 or 30 pounds in different directions, all of the results might be deemed suspect and not useful.

Despite common misconceptions to the contrary, reliability is a property of scores and not of a test or measurement protocol per se.<sup>11</sup> In the recent words of the American Psychological Association Task Force on Statistical Inference,

It is important to remember that a test is not reliable or unreliable. Reliability is a property of the scores on a test for a particular population of examinees . . . Thus, authors should provide reliability coefficients of the scores for the data being analyzed even when the focus of their research is not psychometric.<sup>12</sup>

Put differently, LibQUAL+,<sup>13</sup> like any other measure, is not in and of itself reliable. The psychometric integrity of LibQUAL+<sup>13</sup> scores should be verified every time they are used.

The phase one 2000–2001 academic year version of the measure consists of twenty-five items measuring total service quality, as well as four subdimensions of perceived library quality: (a) *Service Affect* (nine items, such as “willingness to help users”); (b) *Library as Place* (five items, such as “a haven for quiet and solitude”); (c) *Personal Control* (six items, such as “website enabling me to locate information on my own”); and (d) *Information Access* (five items, such as “comprehensive print collections” and “convenient business hours”).

Table 2 presents Cronbach’s alpha coefficients for scores on each of the four dimensions and for total scores. These coefficients indicate a statistically conservative estimate of the proportion of true or meaningful score variance, and are increasingly favorable as they approach 1.0.<sup>13</sup> Analyses are also reported separately for both ARL and non-ARL data, to confirm score reliability as reasonable across institutions of varying types.

**Table 2**  
Score Reliabilities ( $\alpha$ ) Across Scales and Institution Types

Score	Subsample		Total ( <i>n</i> =20,416)
	ARL ( <i>n</i> =16,918)	NonARL ( <i>n</i> =3,498)	
<i>Service Affect</i> (9 items)	.946	.948	.946
<i>Library as Place</i> (5 items)	.928	.929	.929
<i>Personal Control</i> (6 items)	.868	.878	.869
<i>Information Access</i> (5 items)	.758	.766	.758
Total Score (25 items)	.948	.945	.948

### Factor Structure

Factor analysis is a useful technique for determining what dimensions underlie responses to protocols such as LibQUAL+.<sup>13,14</sup> The results also bear upon issues of score validity (i.e., whether scores measure what they are intended to measure). As Nunnally observed, “factor analysis is intimately involved with questions of validity . . . . Factor analysis is at the heart of the measurement of psychological constructs.”<sup>15</sup>

Table 3 presents the pattern/structure coefficients from a principal components analysis of the full data set. These are correlation coefficients between the scores on the twenty-five items and the four factors. Separate analyses were also completed across both institution types (ARL vs nonARL) and random sample splits, and the structure was found not to vary across these divisions. Such replicability lends credence to the validity of the constructs measured across library contexts.

**Table 3**  
**Varimax-rotated Pattern/structure Coefficients**  
 ( $n = 20,416$  Participants;  $v = 25$  items)

No.	Factor				Item Core
	I	II	III	IV	
<i>Service Affect</i>					
18	.85	.13	.24	.13	Willingness to help users
19	.81	.14	.25	.14	Giving users individual attention
4	.81	.18	.13	.21	Employees deal with users caring fashion
25	.79	.19	.19	.12	Employees who are consistently courteous
17	.77	.16	.26	.20	Employees have knowledge answer questions
3	.74	.15	.18	.30	Employees understand needs of users
2	.74	.15	.17	.21	Readiness to respond to users' questions
11	.72	.20	.18	.27	Employees who instill confidence in users
23	.62	.22	.30	.28	Dependability handling service problems
<i>Library as Place</i>					
13	.16	.88	.11	.16	A haven for quiet and solitude
1	.18	.83	.08	.14	A meditative place
12	.22	.83	.15	.22	A contemplative environment
14	.16	.80	.19	.17	Space that facilitates quiet study
22	.20	.80	.18	.20	A place for reflection and creativity
<i>Personal Control</i>					
20	.22	.12	.78	.15	Website enabling me locate info on my own
15	.20	.10	.76	.13	Elec resources accessible home or office
8	.33	.16	.60	.31	Access tools allow me find on my own
24	.30	.24	.59	.29	Modern equip me easily access info I need
10	.35	.18	.56	.42	Info easily accessible for independent use
16	.30	.21	.55	.38	Convenient access to library collections
<i>Information Access</i>					
7	.13	.23	.19	.74	Comprehensive print collections
21	.15	.24	.29	.60	Complete runs of journal titles

No.	Factor				Item Core
	I	II	III	IV	
9	.29	.20	.22	.60	Interdisciplinary library needs addressed
6	.28	.05	.16	.58	Timely document delivery/interlibrary loan
5	.23	.18	.13	.56	Convenient business hours

*Note.* Pattern/structure coefficients greater than |.35| are *italicized*.

## Discussion

Previous analyses have indicated that users think about library service quality using continua rather than in crude discrete categorical terms such as “good” versus “bad”.<sup>16</sup> Other analyses have demonstrated that users tend to take the LibQUAL+™ protocol seriously and do not generate an unacceptable number of illogical or aberrant responses.<sup>17</sup> The present study was conducted to explore score integrity with regard to data from 20,416 participants across a large number of fairly diverse universities.

LibQUAL+™ looks beyond quintessential service issues and includes some items measuring user perceptions of library collections or access—thus the “+” in “LibQUAL+™”. Indeed, failing to ask users at least some questions about collections, even on a measure focusing on service quality, might provoke some puzzlement, because collections do bear on library use experience. However, the measure focuses more on the service dimension than on collection adequacy. Figure 1 presents the dimensions of library service quality that we have seen emerge from the series of studies culminating in the present report.

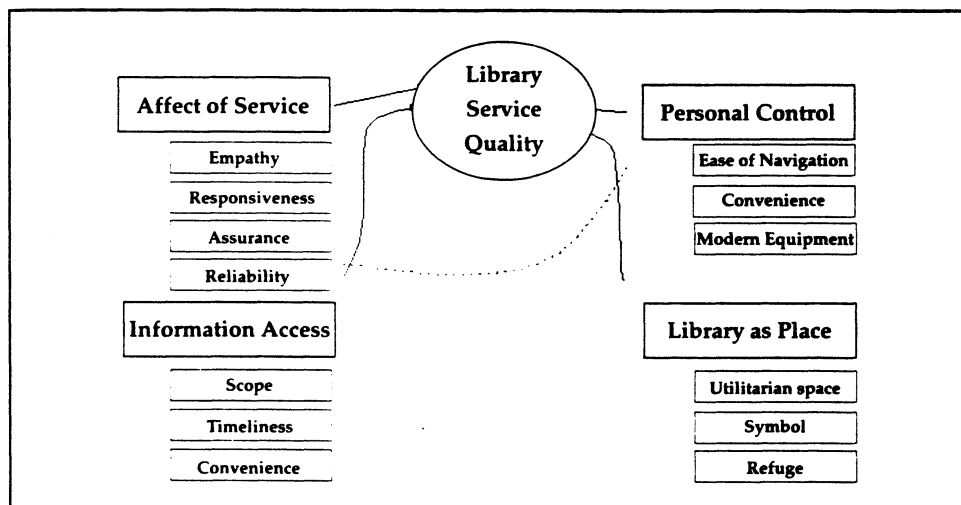


Figure 1. Dimensions of Library Service Quality. © Copyright, Association of Research Libraries (ARL), 2001. Used with permission.

### Score Reliability

Scores from the previous phase zero version of the protocol were highly reliable.<sup>18</sup> Thus, it is not surprising that the scores from the revised and shorter protocol also demonstrate very good reliability, as reported in Table 2.

The scores on the five-item *Information Access* scale were least reliable ( $\alpha = .758$ ). Of course this scale measures some collections-related issues that are not the primary focus of the measure. The scores on the other scales and the total scale are exceptionally reliable.

### Factor Structure

The Table 3 results indicate that the twenty-five items do measure the four intended constructs. These four constructs emerged from the qualitative grounding of the SERVQUAL protocol in the library context.<sup>19</sup> The result is also consistent with previous factor analytic results involving data from 4,407 participants from thirteen ARL institutions with respect to the prior 41-item phase-zero LibQUAL+™ protocol.<sup>20</sup>

### Implications

Large collections do not, in and of themselves, insure that library users always have positive service experiences. Thus, librarians interested in improving service quality need tools to help them benchmark current user perceptions, identify needed areas of improvement, and locate peer institutions obtaining more favorable outcomes. These institutions can then serve as models for improvement programs. LibQUAL+™ is a new and useful tool in this area.

---

**Large collections do not, in and of themselves, insure that library users always have positive service experiences.**

---

LibQUAL+™ should never be the only means that conscientious librarians use to monitor and improve library service quality. Sometimes data will immediately call to mind

particular corrective actions to solve problems. In other instances, focus groups, records of customer complaints or suggestions, or staff reports of user difficulties may be needed to augment LibQUAL+™ data in developing plans of action.

LibQUAL+™, like SERVQUAL,<sup>21</sup> is a total market survey tool. The purpose of a total market survey is to measure customers' or users' overall assessments of service. As Leonard Berry wrote,

When well designed and executed, total market surveys provide a range of information unmatched by any other method . . . A critical facet of total market surveys (and the reason for using the word 'total') is the measurement of competitors' service quality. This requires using noncustomers in the sample to rate the service of their suppliers.<sup>22</sup>

LibQUAL+™ satisfies the major reasons for conducting total market surveys. First, non-local information can reveal how well other libraries perform services and can provide a basis for comparison. Secondly, exemplary libraries can be identified as models for service improvement planning. Finally, total market surveys permit performance tracking over time.<sup>23</sup> Systematic listening to users improves decision making in allocation of scarce resources.



Berry recommends using multiple listening methods and discusses eleven different ways to listen to customers. He writes, "Ongoing data collection through multiple research approaches is a necessity. Transactional surveys, total market surveys, and employee research should always be included."<sup>24</sup> LibQUAL+™ holds promise as a cornerstone in contemporary research libraries' total market survey endeavors.

*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/~bthompson>>.*

*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](mailto:ccook@tamu.edu).*

*Russel L. Thompson is a Systems Analyst at the Texas A&M University Library; he may be contacted via e-mail at: [rltpsych@yahoo.com](mailto:rltpsych@yahoo.com)*

## Notes

This research was supported in part by a grant from the U.S. Department of Education Fund for the Improvement of Postsecondary Education (FIPSE).

1. Danuta A. Nitecki. "Changing the Concept and Measure of Service Quality in Academic Libraries." *The Journal of Academic Librarianship* 22 (1996): 181.
2. Rowena Cullen. "Perspectives on User Satisfaction Surveys." *Library Trends* 49 (2001): 662–63.
3. Nitecki. "Changing the Concept and Measure of Service Quality in Academic Libraries," 181.
4. Colleen Cook, and Fred Heath. "The Association of Research Libraries Libqual+™ Project: An Update." *ARL Newsletter: A Bimonthly Report on Research Library Issues and Actions from ARL, CNI, and SPARC* 211 (2000, August): 12–14; Colleen Cook, Fred Heath, and Bruce Thompson. "LibQUAL+™: One Instrument in the New Measures Toolbox." *ARL Newsletter: A Bimonthly Report on Research Library Issues and Actions from ARL, CNI, and SPARC* 212 (2000, October): 4–7; Colleen Cook, Fred Heath, Bruce Thompson, and Russell L. Thompson. "The Search for New Measures: The ARL LibQUAL+™ Study—A Preliminary Report." *portal: Libraries and the Academy* 1 (2001): 103–12.
5. A. Parasuraman, Leonard L. Berry, and Valarie A. Zeithaml. "Refinement and Reassessment of the SERVQUAL Scale." *Journal of Retailing* 67 (1991): 420–50; A. Parasuraman, Valarie A. Zeithaml, and Leonard L. Berry. "A Conceptual Model of Service Quality and Its Implications for Future Research." *Journal of Marketing* 49 (1985): 41–50; A. Parasuraman, Valarie A. Zeithaml, and Leonard L. Berry. "Alternative Scales for Measuring Service Quality: A Comparative Assessment Based on Psychometric and Diagnostic Criteria." *Journal of Retailing* 70 (1994): 201–30.
6. Valarie A. Zeithaml, A. Parasuraman, and Leonard L. Berry. *Delivering Quality Service: Balancing Customer Perceptions and Expectations*. New York: Free Press, 1990, 16.
7. Colleen Cook, and Bruce Thompson. "Higher-Order Factor Analytic Perspectives on Users' Perceptions of Library Service Quality." *Library and Information Science Research* 22 (2000): 393–404; Colleen Cook, and Bruce Thompson. "Reliability and Validity of SERVQUAL Scores Used to Evaluate Perceptions of Library Service Quality." *Journal of Academic Librarianship* 26 (2000): 248–58; Danuta A. Nitecki. "An Assessment of the Applicability of SERVQUAL Dimensions: A Customer-Based Criteria for Evaluating Quality of Services in an Academic Library," Ph. D. Diss., University of Maryland, 1995. Abstract in *Dissertation Abstracts International* 56:2918A.

8. Colleen Cook and Fred Heath. "Users' Perceptions of Library Service Quality: A LibQUAL+™ Qualitative Interview Study." *Library Trends* 49 (2001): 548–84.
9. Colleen Cook and Bruce Thompson. "Psychometric Properties of Scores from the Web-Based LibQUAL+™ Study of Perceptions of Library Service Quality." *Library Trends* 49 (2001): 585–604.
10. For example, see Cook, Heath, B. Thompson, and R. L. Thompson. "The Search for New Measures," 109.
11. Bruce Thompson. "Guidelines for Authors." *Educational and Psychological Measurement* 54 (1994): 837–47; Bruce Thompson and Tammi Vacha-Haase. "Psychometrics Is Datametrics: The Test Is Not Reliable." *Educational and Psychological Measurement* 60 (2000): 174–95.
12. Leland Wilkinson and the APA Task Force on Statistical Inference. "Statistical Methods in Psychology Journals: Guidelines and Explanations." *American Psychologist* 54 (1999): 594–604. Reprint available through the APA Home Page: <<http://www.apa.org/journals/amp/amp548594.html>> [September 27, 2001].
13. Brian Reinhardt. (1996). "Factors Affecting Coefficient alpha: A Mini Monte Carlo Study." In Bruce Thompson (Ed.), *Advances in social science methodology* 4. Greenwich, CT: JAI Press, 1996: 3–20.
14. Bruce Thompson and Larry G. Daniel. "Factor Analytic Evidence for the Construct Validity of Scores: An Historical Overview and Some Guidelines." *Educational and Psychological Measurement* 56 (1996): 213–24.
15. Jum C. Nunnally. *Psychometric Theory*. 2nd ed. New York: McGraw-Hill, 1978, 112–13.
16. Randall C. Arnau, Russel L. Thompson, and Colleen Cook. "Do Different Response Formats Change the Latent Structure of Responses?: An Empirical Investigation Using Taxonometric Analysis." *Educational and Psychological Measurement* 61 (2001): 23–44.
17. Bruce Thompson, Colleen Cook, and Fred Heath. "The LibQUAL+™ Gap Measurement Model: The Bad, the Ugly, and the Good of Gap Measurement." *Performance Measurement and Metrics* 1 (2000): 165–78.
18. Colleen Cook, Fred Heath, Russel L. Thompson, and Bruce Thompson. "Score Reliability in Web- or Internet-Based Surveys: Unnumbered Graphic Rating Scales Versus Likert-Type Scales." *Educational and Psychological Measurement* 61 (2001): 697–706; B. Thompson, Cook and Heath. "The LibQUAL+™ Gap Measurement Model," 165–178.
19. Cook and Heath. "Users' Perceptions of Library Service Quality." 548–584.
20. Colleen Cook, Fred Heath, and Bruce Thompson. "Users' Hierarchical Perspectives on Library Service Quality: A LibQUAL+™ Study." *College and Research Libraries* 62 (2001): 147–53; Bruce Thompson, Colleen Cook, and Fred Heath. "How Many Dimensions Does It Take to Measure Users' Perceptions of Libraries?: A LibQUAL+™ Study." *portal: Libraries and the Academy* 1 (2001): 129–38.
21. Parasuraman, Berry, and Zeithaml. "Refinement and Reassessment of the SERVQUAL Scale," 420–450.
22. Leonard L. Berry. *On Great Service: A Framework for Action*. New York: The Free Press, 1995, 37.
23. *Ibid.*, 38–9.
24. *Ibid.*, 54.