

How many dimensions does it take to measure users' perceptions of libraries?: A LibQUAL+ study

Bruce Thompson, Colleen Cook and Fred Heath

abstract: The present study is one in a series analyzing the results of a large-scale web-based survey of library service quality at several ARL member institutions. The current study explores the question: what are the fewest dimensions required to measure users' perceptions of library service quality? Understanding the number and hierarchy of dimensions that users employ in evaluating library services is important, because we want to use scoring dimensions that correspond with users' perceptions rather than our own preconceptions of what those perceptions might be. The study was conducted as part of the LibQUAL+ project, which is one of the ARL "New Measures" initiatives. Results for data provided by 3987 participants from eleven ARL institutions were consistent with a view that a single set of scores is one reasonable way to characterize user perceptions of library service quality.

Strong libraries are key components of North American research universities. The library contribution to university research and instruction missions is both important and expensive. The costs of acquiring and disseminating information have accelerated steadily in recent decades. The competition for available resources on every campus has become keen. Without a clearer understanding of outcomes, it is uncertain that the steadily increasing investment in research libraries can be sustained. Without a better understanding of a library's ability to deliver service quality, or where the deficits in delivery may reside, the optimal use of current resources cannot be assured.

It is not surprising therefore, that those libraries confront "increasing pressure... to assess the degree to which their services demonstrate criteria of 'quality.' The emphasis on these measures and services provided to library clientele requires librarians . . . not to equate 'quality' merely with collection size."¹ As Danuta Nitecki noted, "A measure of library quality based solely on collections has become obsolete."² In the contemporary

portal: Libraries and the Academy, Vol. 1, No. 2 (2001), pp. 129-138. Copyright © 2001 by The Johns Hopkins University Press, Baltimore, MD 21218.

Reprinted with permission of The Johns Hopkins University Press.

university context, she argues that "every unit . . . is valued in proportion to its contribution to the quality success of the campus."³

In response to these dynamics, the Association of Research Libraries (ARL) has sponsored a number of "New Measures" initiatives. And, in October 2000, ARL sponsored a two-day conference bringing library directors and staff together with internationally recognized experts on measuring perceptions of service quality. One of the "New Measures" initiatives is the LibQUAL+ study being conducted by ARL and the Texas A&M University Libraries.⁴ Continuing phases of the LibQUAL+ study are being supported in part by the Fund for the Improvement of Post-Secondary Education (FIPSE).

As administered to eleven ARL libraries in spring 2000, LibQUAL+ is a protocol that expands upon and incorporates the twenty-two questions from SERVQUAL, the well-known service quality assessment tool.⁵ SERVQUAL ostensibly measures perceptions and expectations of five dimensions of the general construct of service quality: *tangibles*, *reliability*, *responsiveness*, *assurance*, and *empathy*.⁶ Within this model, "only customers judge quality; all other judgments are essentially irrelevant."⁷

In recent years, the SERVQUAL measure has been administered in a number of library settings. However, it does not appear that the protocol yields the originally intended five dimensions when administered within the library context. Furthermore, as SERVQUAL had its origins in other, mainly for-profit service sectors, it is conceivable that some dimensions of library service that library users deem important are not measured by SERVQUAL.⁸ Therefore, an initial, qualitative research phase of the

LibQUAL+ study involved conducting sixty interviews, at nine ARL institutions, to regroup the twenty-two SERVQUAL items in the library context, and to identify additional items and/or dimensions that may be needed in this particular service environment.⁹

The present study was conducted to explore the question, what are the fewest dimensions required to measure

The present study was conducted to explore the question, what are the fewest dimensions required to measure users' perceptions of library service quality?

users' perceptions of library service quality? However, before addressing this question, a discussion on the its importance may be useful.

Importance of a One-dimensional Perspective

Knowing the number of dimensions that users employ in evaluating library services is important, because it is critical to use scoring dimensions that correspond to users' perceptions rather than librarians' preconceptions. However, it might be desirable from a practical point of view if a reasonable characterization of user perceptions summarized user views in a *single* score. The question, then, is whether it is reasonable to conceptualize a general service quality construct using a single set of scores. This does *not* preclude the possibility that assessing dimensions at various levels of specificity may also be useful.

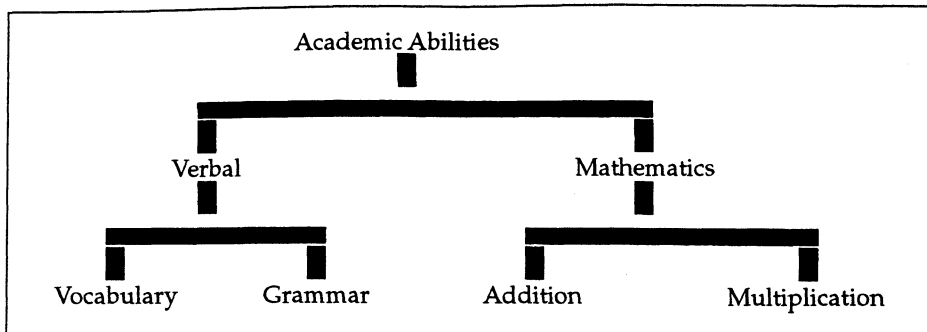


Figure 1. Illustrative Hierarchical (Third-order) Structure

Figure 1 presents an analogy for the measurement of cognitive abilities. At one level, abilities might be conceptualized as four specific knowledge areas or "dimensions" (i.e., Vocabulary, Grammar, Addition, Multiplication). At a slightly higher level of characterization, abilities might be scored as Verbal and Mathematical. Finally, performance might be characterized generically as Academic Abilities.

No one level of characterization has to be employed for all purposes. Indeed, different levels of characterization might be useful for different purposes, and multiple levels of characterization might even be used simultaneously. The same is potentially true in evaluating library users' perceptions of service quality.

Method

Participants

The 3987 participants in the present LibQUAL+ study represented the range of library user groups at eleven ARL member institutions. Included were undergraduate students ($n_u=998$), graduate students ($n_g=1281$), faculty ($n_f=1022$), and staff and ancillary professional personnel (e.g., research scientists not in tenure accruing positions, $n_o=686$).

Half the participants were males, and half females. The ages of the 3987 participants were: (a) < 22 years, 17.8%; (b) 22-30 years, 29.7%; (c) 31-45 years, 25.3%, and (d) > 45 years, 27.2%. The disciplines of the participants were: (a) science, 17.1%; (b) social science, 16.6%; (c) libraries, 12.0%; (d) humanities, 11.3%; (e) engineering, 10.9%; (f) health sciences, 9.4%; (g) business, 8.1%; (h) education, 6.5%; (i) fine arts, 3.8%; (j) law, 1.7%; (k) architecture, 1.5%; and (l) veterinary medicine, 1.0%. Thus, the sample was both large and diverse.

Instrumentation

As noted previously, LibQUAL+ included both the twenty-two items on the SERVQUAL protocol and an additional nineteen items that emerged from qualitative interviews.¹⁰ The forty-one items were administered on the Web using browser software.¹¹ The 3987 participants rated their perceptions of library services using 1-to-9 Likert scales.

Results

Item Analysis from Score Reliability Study

The first analysis explored statistics related to score reliability. As explained in detail by Bruce Thompson and Tammi Vacha-Haase, it is an oxymoron to speak of "the reliability of the test."¹² In the words of the APA 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.¹³

To evaluate the psychometric integrity of the scores from these 3987 participants, we computed a number of the conventional item-analysis statistics reported in table 1. First, we computed so-called "corrected" item discrimination coefficients. These are correlations between item scores (i.e., 1 to 9) with scores on all remaining forty items excluding the given item under consideration (here ranging from 40 [40 x 1] to 360 [40 x 9]). If all items are homogeneous in their content, these coefficients will tend to be positive and large.

Second, we computed R^2 values between scores on a given item and the scores on the remaining forty items. Third, we computed a-if-deleted statistics for the forty-one items. These are the alpha coefficients for scores on forty items when a given item was deleted.

The alpha coefficient for scores on the forty-one items was .965. This is a strikingly large value, and strongly indicates homogeneity of item content or high item score covariance. However, it is important not to interpret this remarkable value as meaning that the users' perceptions can *only* be characterized with a single score. Conceptually discrete subscales on a measure may have quite high alpha coefficients, but still aggregate at high orders of analysis into overarching constructs.¹⁴

Factor Analytic Results

We conducted a principal components analysis in which a single factor was extracted to explore the structure underlying the scores.¹⁵ The factor pattern/structure coefficients from this analysis are presented in table 2. Also reported in the table are coefficients computed independently for the three larger user groups: undergraduate and graduate students and faculty.

Discussion

All the table 1 results are consistent with a view that a single set of scores is one reasonable way to characterize user perceptions of library service quality. The "corrected" item discrimination coefficients ranged from .51 to .75. The R^2 values of scores on a given item with scores on the remaining forty items ranged from .32 to .76. The alpha coefficient for the forty-one items was .965, and all a-if-item-deleted statistics remained quite large.

Table 1
Item Statistics from Reliability Analyses

Item Core/Topic	"Corrected" Discrimination Coefficient	Squared Multiple Correlation	α if Item Deleted
<i>Tangibles</i>			
6 Visually appealing facilities	.60	.56	.9641
23 Visual appeal materials	.67	.51	.9637
35 Modern equipment	.61	.42	.9640
38 Employees have neat appearance	.56	.38	.9642
<i>Reliability</i>			
2 Providing services as promised	.64	.59	.9639
5 Service at promised time	.66	.59	.9638
16 Maintain error-free records	.60	.42	.9641
28 Performing services right	.72	.64	.9636
33 Handle users' service problems	.70	.62	.9637
<i>Responsiveness</i>			
3 Keep users informed	.56	.43	.9642
18 Readiness to respond to user	.70	.74	.9636
19 Willingness to help users	.69	.76	.9636
26 Prompt service to users	.75	.65	.9635
<i>Assurance</i>			
9 Employees instill confidence	.68	.61	.9637
20 Employees have knowledge	.70	.67	.9636
31 Assuring accuracy/confidentiality	.59	.41	.9641
34 Employees who are courteous	.68	.64	.9637
<i>Empathy</i>			
8 Convenient business hours	.52	.32	.9644
11 Employees understand needs	.73	.68	.9634
13 Users' best interests at heart	.73	.60	.9634
24 Deal with users in caring fashion	.73	.69	.9635
41 Giving users individual attention	.69	.57	.9636
<i>Access to Collections</i>			
1 Convenient access collections	.60	.46	.9641
7 Access to archives	.59	.41	.9641
10 Resources added to collection	.56	.40	.9642
15 Instruction in use, when needed	.64	.49	.9639
17 Timely document delivery	.56	.40	.9642
25 Full-text delivered electronically	.51	.32	.9645
27 Comprehensive print collection	.60	.48	.9640

Table 1 Cont.

Item Core/Topic	"Corrected" Discrimination Coefficient	Squared Multiple Correlation	α if Item Deleted
32 Library materials in the stacks	.57	.39	.9642
36 Interdisciplinary needs addressed	.59	.43	.9641
37 Complete runs of journal titles	.56	.47	.9642
<i>Library as Place</i>			
4 A place for reflection	.57	.55	.9643
12 A contemplative environment	.65	.68	.9638
14 Comfortable and inviting location	.67	.65	.9637
21 A secure and safe place	.55	.34	.9643
22 Center intellectual interaction	.58	.50	.9642
29 Space group/individual study	.59	.50	.9641
30 A haven for quiet and solitude	.59	.68	.9642
39 A meditative place	.58	.67	.9642
40 Space that facilitates quiet	.60	.70	.9640

Similarly, the factor pattern/structure coefficients for the 3987 participants ranged from .531 to .780, as reported in table 2. It is also striking that the one-dimensional structures computed separately for three of the user groups were so similar. Thus undergraduates, graduate students, and faculty seem to hold a similar mental construct of service quality in academic libraries.

These results suggest that users' perceptions of library service quality can be characterized using a single overarching dimension; i.e., service quality. However, it is noteworthy that this general construct is saturated with elements of service based on users' perceptions that library staff understand and care about users' service needs. For example, as reported in table 2, the four items most correlated with the single service factor were respectively:

- 26 Prompt service to users (.780);
- 11 Employees understand needs of their users (.769);
- 24 Employees who deal with users in caring fashion (.764);
- 13 Having the users' best interest at heart (.760).

As Colleen Cook and Russell L. Thompson explained, these findings do not mean that users ignore library collections or the library as place in their thinking.¹⁶ However, these results do indicate that service quality perceptions (e.g., empathy, reliability, responsiveness) are especially important in defining the dimensions of users' views of libraries. Thus, the LibQUAL+ and other ARL "New Measures" initiatives measure issues of considerable interest to library users.

Table 2

Factor Pattern/Structure Coefficients for One-Factor Solutions

Scale	No.	Group			
		Ugrads (n=998)	Grads (n=1281)	Faculty (n=1022)	Total (n=3987)
Responsiveness	26	.78	.79	.79	.780
Empathy	11	.77	.77	.78	.769
Empathy	24	.76	.76	.78	.764
Empathy	13	.77	.76	.76	.760
Reliability	28	.75	.74	.78	.755
Assurance	20	.75	.73	.76	.744
Responsiveness	18	.78	.74	.76	.742
Reliability	33	.73	.72	.79	.740
Responsiveness	19	.78	.72	.76	.739
Empathy	41	.70	.74	.77	.727
Assurance	34	.74	.70	.75	.717
Assurance	9	.72	.71	.74	.714
Tangibles	23	.69	.66	.71	.688
Reliability	5	.65	.70	.72	.687
Collections	15	.69	.65	.71	.676
Reliability	2	.64	.68	.70	.671
Place	14	.67	.66	.68	.671
Place	12	.69	.66	.62	.646
Tangibles	35	.64	.61	.67	.628
Reliability	16	.65	.61	.66	.627
Collections	27	.65	.63	.62	.622
Collections	1	.63	.62	.64	.620
Assurance	31	.61	.62	.65	.616
Collections	7	.64	.62	.64	.612
Collections	36	.65	.61	.60	.612
Tangibles	6	.58	.60	.62	.608
Place	40	.57	.63	.61	.603
Collections	32	.63	.64	.57	.599
Responsiveness	3	.54	.61	.65	.595
Collections	17	.60	.59	.63	.592
Place	30	.58	.58	.62	.590
Tangibles	38	.60	.60	.59	.589
Place	29	.58	.59	.61	.589
Collections	10	.58	.61	.59	.584
Place	22	.62	.58	.56	.583

Table 2 Cont.

Scale	No.	Group			Total (n=3987)
		Ugrads (n=998)	Grads (n=1281)	Faculty (n=1022)	
Collections	37	.60	.60	.57	.582
Place	4	.55	.57	.58	.571
Place	21	.63	.56	.58	.569
Empathy	8	.54	.53	.58	.545
Collections	25	.55	.52	.52	.531

Having established that a single dimension can express user perceptions of library service quality, the next questions from the library community can be anticipated. If there is a single dimension, what is its score? Further, how do I interpret that score to place my library in context? Can my library be ranked against others in order to place perceptions of quality in the context of my peers?

The answer to which single score best serves to define the dimension merits further psychometric evaluation, and will be the focus of further study. Inevitably, a single factor score raises the issue of ranking. While ranking libraries on the basis of scores is theoretically sound, we urge the cautious exploration of that path. While the elements of service measured by LibQUAL+ would seem to cut across all libraries, the thoughtful administrator would exercise caution in making generalizations. It would be unwise to take the entire result set, place the libraries in rank order, and—in isolation from institution-specific contexts offer up that list as the definitive measure of service quality in academic libraries.

As library directors consider best practice options, the roles and missions of other university research libraries should always be an important consideration. Directors of land-grant universities might look to others in that cohort group to find libraries with practices worthy of study or emulation.

As library directors consider best practice options, the roles and missions of other university research libraries should always be an important consideration.

Private universities with comprehensive roles might find it most useful to look within that sub-group for libraries with high scores for their models. Technical universities with highly focused missions might also usefully study each other's results. Just as we know that IQ scores, as widely accepted and useful measures, are not the singular, definitive

metrics of all intelligence components, a single score for library service quality may be a useful evaluative tool, but one that should be applied in full understanding of its limitations. The next round of the LibQUAL+ pilot studies should provide additional information to help address the issue of institutional rankings.

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://acs.tamu.edu/~bbt6147/>.

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 address: ccook@tamu.edu.

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 address: fheath@tamu.edu.

Notes

1. Peter Herson and Charles R. McClure, *Evaluation and Library Decision-Making*, Norwood, New Jersey: Ablex, 1990, xv.
2. Danuta A. Nitecki, "Changing the Concept and Measure of Service Quality in Academic Libraries," *Journal of Academic Librarianship* 22 (1996): 181.
3. *Ibid.*, 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): 12-14.
5. Colleen Cook, Fred Heath, Bruce Thompson, and Russel L. Thompson, "The Search for New Measures: The ARL 'LibQUAL+' Study—A Preliminary Report," *Portal: Libraries and the Academy* 1.1 (2001).
6. See A. Parasuraman, Leonard L. Berry and Valarie A. Zeithaml, "SERVQUAL: A Multiple-item Scale for Measuring Consumer Perceptions of Service Quality," *Journal of Retailing* 64 (1988): 12-40; A. Parasuraman, Leonard L. Berry and Valarie A. Zeithaml, "Refinement and Reassessment of the SERVQUAL Scale," *Journal of Retailing* 67 (1991): 420-450; 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): 201-230.
7. Valarie A. Zeithaml, A. Parasuraman and Leonard L. Berry, *Delivering Quality Service: Balancing Customer Perceptions and Expectations*, New York: Free Press (1990), 16.
8. Colleen Cook and Bruce Thompson, "Higher-order Factor Analytic Perspectives on Users' Perceptions of Library Service Quality," *Library Information Science Research* (in press); 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): 248-258; 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.
9. Colleen Cook and Fred Heath, "Users' Perceptions of Library Service Quality: A 'LibQUAL+' Qualitative Interview Study," *Library Trends* (in press). [Paper presented at the Association of Research Libraries Measuring Service Quality Symposium on the New Culture of Assessment: Measuring Service Quality, Washington, D.C., October 20, 2000].
10. Colleen Cook and Fred Heath, "The Association of Research Libraries LibQUAL+ Project: An Update," 12-14.
11. Colleen Cook, Fred Heath, Russel L. Thompson, and Bruce Thompson, "Score Reliabilities in Web- or Internet-based Surveys: Unnumbered Graphic Rating Scales Versus Likert

- Scales," *Educational and Psychological Measurement* (in press). [Paper presented at the Association of Research Libraries Measuring Service Quality Symposium on the New Culture of Assessment: Measuring Service Quality, Washington, D.C., October 20, 2000.]
12. Bruce Thompson and Tammi Vacha-Haase, "Psychometrics is Datametrics: The Test is Not Reliable," *Educational and Psychological Measurement* 60 (2000): 174-195.
 13. Leland Wilkinson and The APA Task Force on Statistical Inference, "Statistical Methods in Psychology Journals: Guidelines and Explanations," *American Psychologist* 54 (1999): 594-604.
 14. Linz Crocker and James Algina, *Introduction to Classical and Modern Test Theory*, New York: Holt, Rinehart and Winston, 1986.
 15. Roderick D. Hertzler, "A Primer on Factor Analysis with Comments on Analysis and Interpretation Patterns," in Vol. 4 of *Advances in Social Science Methodology*, ed. Bruce Thompson, (Greenwich, CT: JAI Press), 175-206. 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): 213-224.
 16. Colleen Cook and Bruce Thompson, "Psychometric Properties of Scores from Web-based LibQUAL+ Study of Perceptions of Library Service Quality."