

Scaling for the LibQUAL+™ Instrument: A Comparison of
Desired, Perceived and Minimum Expectation Responses
versus Perceived Only

Colleen Cook

Texas A&M University

Bruce Thompson

Texas A&M University

and

Baylor College of Medicine

Paper presented at the 4th Northumbria International Conference, Pittsburgh, Pennsylvania, USA, August 14, 2001. This research was supported in part by a grant from the U.S. Department of Education Fund for the Improvement of Postsecondary Education (FIPSE).

Colleen Cook may be contacted via e-mail at address: "ccook@tamu.edu". Bruce Thompson may be contacted via Internet URL: "http://www.coe.tamu.edu/~bthompson".

A periodically-updated bibliography of LibQUAL+™ studies may be accessed at: "http://www.coe.tamu.edu/~bthompson/servqbib.htm".

<<http://www.coe.tamu.edu/~bthompson/libqual5.htm>>

Abstract

The LibQUAL+™ web-based protocol was recently completed by 20,416 participants representing 43 universities. The present study investigated whether the perceptions of participants randomly assigned the LibQUAL+™ short form (target: 40% of respondents) were comparable to those of participants responding in all three response frameworks (minimally-acceptable service quality; perceived levels of service quality; and desired levels of service quality).

Those who use, fund, and work in contemporary research libraries increasingly recognize that "A measure of library quality based solely on collections has become obsolete" (Nitecki, 1996, p. 181). This realization led the Association of Research Libraries (ARL), whose membership includes the largest 123 research libraries in North America, to initiate the "New Measures" program. One of these initiatives is the LibQUAL+™ research and development project (Cook & Heath, 2000; Cook, Heath & B. Thompson, 2000).

The LibQUAL+™ research and development project is an ARL/Texas A&M University joint effort. This project is also supported, in part, by a multi-year grant from the U.S. Department of Education's Fund for the Improvement of Post-Secondary Education (FIPSE).

Academic Year 1999-2000

During the 1999-2000 academic year LibQUAL+™ was completed on the Web 4,407 by participants from 13 ARL institutions. This form of the protocol included 22 items from the well-known SERVQUAL measure (cf. Parasuraman, Berry & Zeithaml, 1991; Parasuraman, Zeithaml & Berry, 1985, 1994). The SERVQUAL protocol has been used previously in library service quality research (cf. Cook & Thompson, 2000a, 2000b). Participants also completed 19 trial items that were developed, following qualitative analysis of interviews of library users at nine universities (Cook & Heath, 2001), to measure service quality features unique to the library setting.

A series of articles have been published to report analyses of these 1999-2000 data (Arnau, R.L. Thompson & Cook; 2001; Cook, Heath & B. Thompson, 2001; Cook, Heath, B. Thompson & R.L. Thompson, 2001; Cook, Heath, R.L. Thompson & B. Thompson, 2001; Cook & B. Thompson, 2001; B. Thompson, 2000; B. Thompson, Cook & Heath, 2000; B. Thompson, Cook & Heath, 2001). Following these analyses, protocol items were further revised.

Academic Year 2000-2001

In the spring of 2001, the items on the web-administered LibQUAL+™ protocol were completed by 20,416 participants from 43 campuses. Of these 43 libraries, 35 were members of ARL.

A series of reports associated with the final LibQUAL+™ measure have been written (cf. Cook, Heath & B. Thompson, 2001; Heath, Cook, Kyrillidou & B. Thompson, 2001; B. Thompson, Cook & R.L. Thompson, 2001). These various analyses have indicated that:

--the 25-item LibQUAL+™ survey yields reliable scores on 4 scales (i.e., Service Affect, Library as Place, Personal Control, and Information Access) as well as on the total scale;

--the factor structure underlying responses matches the expected structure;

--both individual and institutional normative tables for converting scale and total scores into standardized scores and percentile rank scores can (and have) been developed; and

--scale and total scores correlate highly with perceptions of service quality, but not with collections count measures, such as ARL Index scores, as expected.

Two Interpretation Frameworks

The 25 LibQUAL+™ items, the four subscales, and the total score are all measured on a 1-to-9 scale, with higher scores representing more positive user perceptions. Let's assume that a given library received a total LibQUAL+™ score mean of 6.55. This mean is higher than the scale midpoint of 5.0.

This seems to be a positive result, given the scale midpoint. But is this result really positive, and if so, how positive? There are two interpretation strategies that are both considerably more informative than interpreting the mean of 6.55 against the reference of the scale midpoint of 5.0.

Score Norms. Preliminary score norms have been developed for LibQUAL+™ (cf. Cook, Heath & B. Thompson, 2001). Norm conversion tables facilitate the interpretation of observed scores using norms created for a large and representative sample. The LibQUAL+™ norm conversion tables have been created at both the individual and the institutional levels.

For example, individual norm tables might indicate that only 49% of library users rate their libraries 6.55 or lower. However, institutional norm tables might indicate that only 27% of all university libraries in the normative sample received a mean of 6.55 or lower. The example makes clear that interpretations of results using norms versus response scale midpoints may differ quite markedly!

Zones of Tolerance. The SERVQUAL protocol asks participants to rate service quality based on (a) minimally-acceptable service quality; (b) perceived levels of service quality; and (c) desired levels of service quality. The mean ratings of minimally-acceptable and desired service quality ratings create a "zone of tolerance" for customer perceptions (cf. Parasuraman, Zeithaml & Berry, 1985, 1994). It is generally hoped that perceived service-quality ratings will fall within the zone, and that in particular perceived quality will not be below minimally-acceptable service quality.

The "zone of tolerance" interpretative framework has intuitive appeal (Thompson, Cook & Heath, 2000). However, it must be remembered that interpretations in this framework may differ from those made in the normative interpretation framework. For example, library users seem to be uniformly dissatisfied with runs of journal titles, and thus mean perceived ratings of this item tend routinely to be lower than minimally-acceptable

ratings (cf. Cook, 2001). Only a normative interpretation reveals this sort of dynamic by framing scores within a reference group profile.

Purpose of the Present Study

A critical consideration in surveying is representativeness; response rates are also relevant to the extent that response rates impact representativeness (Thompson, 2000). For this reason survey researchers strive to maximize return. A comprehensive meta-analysis of factors impacting participation in web-based surveys indicates that survey length bears upon response rate (Cook, Heath & R.L. Thompson, 2000).

Web-based surveys have tremendous flexibility. An example of such flexibility with regard to LibQUAL+TM involved the potential to determine randomly whether respondents were asked to rate library service quality using all three response frameworks (i.e., minimally-acceptable, perceived, desired) or only perceptions of library service quality. Thus the boundaries of the interpretive zones of tolerance using only a subsample of responses could be benchmarked, while gathering complete data on perceptions, which are the primary focus of the evaluation.

The present study investigated whether the perceptions of participants receiving the short form (target: 40% of respondents) were comparable to those of participants responding in all three response frameworks (target: 60% of respondents). The analysis focused on the final pool of 25 LibQUAL+TM items that delineate four subscales (Service Affect, Library as Place, Personal Control, and Information Access) and total scores (cf. Cook, Heath & B. Thompson, 2001; Heath, Cook, Kyriellidou & B. Thompson, 2001; B. Thompson, Cook & R.L. Thompson, 2001).

Theoretically, if (a) response framework is randomly assigned and (b) the response format does not itself impact perceptions, then perceptions should be comparable across both frameworks. In the presence of random assignment to response format groups, such a finding might suggest therefore that the boundaries of the zones of tolerance may generalize across both groups.

Results

Table 1 presents the means and standard deviations of the four LibQUAL+TM subscales and total scores across forms. To evaluate result replicability, which is not addressed by statistical significance testing (Cohen, 1994; Thompson, 1996), separate analyses were computed across ARL and nonARL participants.

The table also presents 95% confidence intervals about means. Confidence intervals are not yet widely used and are poorly understood (Fidler & Thompson, 2001; Finch, Cumming & Thomason, 2001). Nevertheless, confidence intervals are very useful statistical tools (Wilkinson & APA Task Force on Statistical Inference, 1999). As the American Psychological Association recently noted,

The reporting of confidence intervals... can be an extremely effective way of reporting results. Because confidence intervals combine information on location and precision and can often be directly used to infer significance levels, they are, in general, the best reporting strategy. The use of confidence intervals is therefore strongly recommended. (p. 22, emphasis added)

Table 1 also presents the eta squared (η^2) effect size that characterizes the percentage of variability of LibQUAL+TM scores that can be predicted from knowledge of whether participants received the short or the long form. As the American Psychological Association recently noted,

For the reader to fully understand the importance of your findings, it is almost always necessary to include some index of effect size or strength of relationship in your Results section.... The general principle to be followed... is to provide the reader not only with information about statistical significance but also with enough information to assess the magnitude of the observed effect or relationship. (pp. 25-26, emphasis added)

Discussion

The means presented in Table 1 were consistently smaller for short form participants across scales and respondent groups. Furthermore, score standard deviations were larger for short form participants; long form participants gave responses that were somewhat less variable. However, the magnitudes of these differences were very small. All the eta squared (η^2) effect sizes (analogous to r^2) were less than 1%.

In the aggregate these results suggest that the protocol of randomly assigning the short form measuring only perceptions to some respondents can minimize response burdens on the participant pool while still providing reasonable estimates of the zones of tolerance boundaries. Of course, these results remain to be replicated during LibQUAL+TM phase two (2001-2002).

During the academic year 2001-2002, we anticipate that library users from roughly 200 institutions will participate in the LibQUAL+TM survey. Further information about the project may be accessed via URL: "<http://www.arl.org/libqual>".

References

American Psychological Association. (2001). Publication manual of the American Psychological Association (5th ed.). Washington, DC: Author.

Arnau, R.C., Thompson, R.L., & Cook, C. (2001). Do different response formats change the latent structure of responses?: An empirical investigation using

- taxonometric analysis. *Educational and Psychological Measurement*, 61, 23-44.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997-1003.
- Cook, C.C. (2001). A mixed-methods approach to the identification and measurement of academic library service quality constructs: LibQUAL+TM . (Doctoral dissertation, Texas A&M University). ??
- Cook, C., & Heath, F. (2000, August). The Association of Research Libraries LibQUAL+TM Project: An update. *ARL Newsletter: A Bimonthly Report on Research Library Issues and Actions from ARL, CNI, and SPARC*, 211, 12-14.
- Cook, C., & Heath, F. (2001). Users' perceptions of library service quality: A "LibQUAL+TM " qualitative interview study. *Library Trends*, 49(4), ??.
- Cook, C., Heath, F., & Thompson, B. (2000, October). LibQUAL+TM : One instrument in the New Measures toolbox. *ARL Newsletter: A Bimonthly Report on Research Library Issues and Actions from ARL, CNI, and SPARC*, 212, 4-7.
- Cook, C., Heath, F., & Thompson, B. (2001). Users' hierarchical perspectives on library service quality: A "LibQUAL+TM " study. *College and Research Libraries*, 62, 147-153.
- Cook, C., Heath, F. & Thompson, B. (in press). Score norms for improving library service quality: A LibQUAL+TM study. *Portal: Libraries and the Academy*. [ms submitted]
- Cook, C., Heath, F., Thompson, B., & Thompson, R.L. (2001). The search for new measures: The ARL "LibQUAL+TM " study--a preliminary report. *Portal: Libraries and the Academy*, 1, 103-112.
- Cook, C., Heath, F., & Thompson, R.L. (2000). A meta-analysis of response rates in Web- or Internet-based surveys. *Educational and Psychological Measurement*, 60, 821-836.
- Cook, C., Heath, F., Thompson, R.L. & Thompson, B. (2001). Score reliability in Web- or Internet-based surveys: Unnumbered graphic rating scales versus Likert-type scales. *Educational and Psychological Measurement*, 61, 697-706.
- Cook, C., & Thompson, B. (2000a). Higher-order factor analytic perspectives on users' perceptions of library service quality. *Library Information Science Research*, 22, 393-404.
- Cook, C., & Thompson, B. (2000b). Reliability and validity of SERVQUAL scores used to evaluate perceptions of library service quality. *Journal of Academic Librarianship*, 26, 248-258.

Cook, C., & Thompson, B. (2001). Psychometric properties of scores from the Web-based LibQUAL+™ study of perceptions of library service quality. *Library Trends*, 49(4), ??.

Fidler, F., & Thompson, B. (2001). Computing correct confidence intervals for ANOVA fixed- and random-effects effect sizes. *Educational and Psychological Measurement*, 61, 575-605.

Finch, S., Cumming, G., & Thomason, N. (2001). Reporting of statistical inference in the *Journal of Applied Psychology*: Little evidence of reform. *Educational and Psychological Measurement*, 61, 181-210.

Heath, F., Cook, C., Kyrillidou, M., & Thompson, B. (in press). ARL Index and other validity correlates of LibQUAL+™ scores. *Portal: Libraries and the Academy*. [ms submitted]

Nitecki, D.A. (1996). Changing the concept and measure of service quality in academic libraries. *The Journal of Academic Librarianship*, 22, 181-190.

Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher*, 25(2), 26-30.

Thompson, B. (2000, October). Representativeness versus response rate: It ain't the response rate!. Paper presented at the Association of Research Libraries (ARL) Measuring Service Quality Symposium on the New Culture of Assessment: Measuring Service Quality, Washington, DC.

Thompson, B., Cook, C., & Heath, F. (2000). The LibQUAL+™ gap measurement model: The bad, the ugly, and the good of gap measurement. *Performance Measurement and Metrics*, 1, 165-178.

Thompson, B., Cook, C., & Heath, F. (2001). How many dimensions does it take to measure users' perceptions of libraries?: A "LibQUAL+™" study. *Portal: Libraries and the Academy*, 1, 129-138.

Thompson, B., Cook, C., & Thompson, R.L. (in press). Reliability and structure of LibQUAL+™ scores. *Portal: Libraries and the Academy*. [ms submitted]

Wilkinson, L., & APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594-604. [reprint available through the APA Home Page: <http://www.apa.org/journals/amp/amp548594.html>]

Table 1

ANOVAs of LibQUAL+TM T Scores by Form

Group	Count	Mean (SD)	95% CI for Mean	eta sq
ARL Participants (n = 16,918)				
Service Affect				
Short	8344	49.6311(10.6428)	49.4027 to 49.8595	0.13%
Long	8574	50.3590(9.3186)	50.1618 to 50.5563	
Total	16918	50.0000(10.0000)	49.8493 to 50.1507	
Fcalc = 22.44; df = 1/16,916; pcalc = .000002				
Library as Place				
Short	8344	49.5637(10.8562)	49.3307 to 49.7966	0.18%
Long	8574	50.4246(9.0702)	50.2326 to 50.6167	
Total	16918	50.0000(10.0000)	49.8493 to 50.1507	
Fcalc = 31.40; df = 1/16,916; pcalc = .00000002				
Personal Control				
Short	8344	49.7153(10.6931)	49.4859 to 49.9448	0.08%
Long	8574	50.2770(9.2682)	50.0808 to 50.4732	
Total	16918	50.0000(10.0000)	49.8493 to 50.1507	
Fcalc = 13.35; df = 1/16,916; pcalc = .0002				
Information Access				
Short	8344	49.6638(10.6960)	49.4343 to 49.8934	0.11%
Long	8574	50.3272(9.2616)	50.1311 to 50.5232	
Total	16918	50.0000(10.0000)	49.8493 to 50.1507	
Fcalc = 18.62; df = 1/16,916; pcalc = .000002				
LibQUAL+ β Total Score				
Short	8344	49.5634(10.5989)	49.3360 to 49.7909	0.18%
Long	8574	50.4248(9.3616)	50.2267 to 50.6230	
Total	16918	50.0000(10.0000)	49.8493 to 50.1507	
Fcalc = 31.43; df = 1/16,916; pcalc = .000000002				
nonARL Participants (n = 3,498)				
Service Affect				
Short	1672	49.6365(10.3966)	49.1378 to 50.1352	0.12%

Long 1826 50.3329(9.6133) 49.8916 to 50.7741
Total 3498 50.0000(10.0000) 49.6685 to 50.3315
Fcalc = 4.24; df = 1/3,496; pcalc = .040

Library as Place

Short 1672 49.2130(10.7440) 48.6976 to 49.7283 0.57%
Long 1826 50.7207(9.2107) 50.2979 to 51.1434
Total 3498 50.0000(10.0000) 49.6685 to 50.3315
Fcalc = 19.95; df = 1/3,496; pcalc = .0000009

Personal Control

Short 1672 49.7500(10.2555) 49.2580 to 50.2419 0.06%
Long 1826 50.2289(9.7574) 49.7811 to 50.6768
Total 3498 50.0000(10.0000) 49.6685 to 50.3315
Fcalc = 2.00; df = 1/3,496; pcalc = .157

Information Access

Short 1672 49.5016(10.5283) 48.9966 to 50.0066 0.23%
Long 1826 50.4563(9.4705) 50.0217 to 50.8910
Total 3498 50.0000(10.0000) 49.6685 to 50.3315
Fcalc = 7.97; df = 1/3,496; pcalc = .005

LibQUAL+ Total Score

Short 1672 49.4375(10.1910) 48.9487 to 49.9264 0.29%
Long 1826 50.5150(9.7964) 50.0654 to 50.9646
Total 3498 50.0000(10.0000) 49.6685 to 50.3315
Fcalc = 10.16; df = 1/3,496; pcalc = .001
