Item sampling in service quality assessment surveys to improve response rates and reduce respondent burden
The "LibQUAL + ® Lite" example

Bruce Thompson
Texas A&M University and Baylor College of Medicine, College Station, Texas, USA

Martha Kyrillidou
Association of Research Libraries, Washington, DC, USA, and

Colleen Cook
Texas A&M University, College Station, Texas, USA

Abstract

Purpose – Survey researchers sometimes develop large pools of items about which they seek participants' views. As a general proposition, library participants cannot reasonably be expected to respond to 100 + items on a given service quality assessment protocol. This paper seeks to describe the use of matrix sampling to reduce that burden on the participant.

Design/methodology/approach – Matrix sampling is a survey method that can be used to collect data on all survey items without requiring every participant to react to every survey question. Here the features of data are investigated from one such survey, the LibQUAL + ® Lite protocol, and the participation rates, completion times, and result comparisons across the two administration protocols – the traditional LibQUAL + ® protocol and the LibQUAL + ® Lite protocol – at each of the four institutions are explored.

Findings – Greater completion rates were realized with the LibQUAL + ® Lite protocol.

Originality/value – The data from the Lite protocol might be the most accurate representation of the views of all the library users in a given community.

Keywords Information services, Service quality assurance, Quality assessment, Surveys

Paper type Research paper

Introduction

As Rowena Cullen noted, "focusing more energy on meeting [...] customers' expectations" is critical in the contemporary academic library environment, in part because "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" (Cullen, 2001).

In this environment, as Danuta Nitecki has observed, "A measure of library quality based solely on collections [counts] has become obsolete" (Nitecki, 1996). Librarians have come to realize the wisdom of the words of French philosopher and moralist François de La Rochefoucauld: "Il est plus nécessaire d'étudier les hommes que les

livres" (de La Rochefoucauld, 1613-1680). In the words of Bruce Thompson, "We only care about the things we measure" (Thompson, 2006), so we do not seriously care about service quality unless we listen to library users in various systematic ways. Within a service quality orientation, "only customers judge quality; all other judgments are *essentially irrelevant*" (Zeithaml *et al.*, 1990).

*LibQUAL + *®

One service quality assessment tool that has been widely used in libraries around the world is LibQUAL + ®. LibQUAL + ® has three primary components. As noted elsewhere:

> First, LibQUAL + ® consists of 22 core items measuring perceived service quality with respect to (a) Service Affect, (b) Library as Place, and (c) Information Control. Each item is rated with respect to (a) minimally-acceptable service expectations, (b) desired service expectations, and (c) perceived level of actual service quality [...] Second, the LibQUAL + ® protocol solicits open-ended comments from users regarding library service quality [...] These comments are crucial, because here the participants elaborate upon perceived strengths and weaknesses, and sometimes offer suggestions for specific actions to improve service. Third, libraries using LibQUAL + ® have the option of selecting five additional items from a supplementary pool of 100 + items to augment the 22 core items to focus on issues of local interest (Thompson *et al.*, 2007).

LibQUAL + ® data can be evaluated using any combination of three interpretation frameworks:

(1) location of perceptions within the "zones of tolerance" defined by minimally acceptable and desired expectations;

(2) benchmarking against peer institutions; and

(3) comparing changes in a given institution's data longitudinally over time.

In the ten years since its inception in 2000 (Thompson, 2007), LibQUAL + ® has been used to collect data from more than 1.25 million library users from more than 1,000 institutions! LibQUAL + ® has now been used in 22 different countries: the USA, Canada, Mexico, Bahamas, Australia, New Zealand, the UK (England, Scotland, Wales), France, Ireland, Belgium, The Netherlands, Switzerland, Denmark, Finland, Norway, Sweden, Egypt, United Arab Emirates, South Africa, Hong Kong, Singapore, and Japan. Currently, the system supports 15 languages: Afrikaans, American English, British English, Chinese (Traditional), Danish, Dutch, Finnish, French (Canadian), French (European), German, Norwegian, Spanish, Swedish, Welsh and Japanese. The development and use of LibQUAL + ® has been documented in a host of academic outlets (Cook *et al.*, 2001c; Thompson *et al.*, 2001 Cook *et al.*, 2003; Heath *et al.*, 2002; Cook *et al.*, 2002; Cook and Heath, 2001; Cook *et al.*, 2001a; Thompson *et al.*, 2005, 2005, 2007, 2008.

Purposes of the present article

Survey researchers sometimes develop large pools of items about which they seek participants' views. For example, in the Association of Research Libraries DigiQUAL® (Association of Research Libraries) project, the item pool consist of more than 100 items. As a general proposition, library participants cannot reasonably be expected to respond to 100 + items on a given service quality assessment protocol. However, a

survey method called "matrix sampling" can be used to collect data on all survey items without requiring every participant to react to every survey question. Here we investigate the features of data from one such survey – the LibQUAL + [®] Lite protocol.

LibQUAL + [®] Lite is a survey methodology in which:

- *all* users answer a few, selected survey questions (i.e. three core items); but
- the remaining survey questions are answered *only* by a randomly selected subsample of the users.

Thus, data are collected on *all* questions, but each user answers *fewer* questions, thus shortening the required response time. Table I illustrates this survey strategy. In this example, all users complete three of the items (i.e. the first, second, and fourth items). But only Mary and Sue were randomly selected to complete the third item in the item pool, which was Service Affect item #2. Only Bob and Mary were randomly selected to complete the fifth item in the item pool, which was Service Affect item #3. Only Sue and Ted were randomly selected to complete the sixth item in the item pool, which was Information Control item #2.

On LibQUAL + [®] Lite, each participant completes only eight of the 22 core survey items. Every participant completes the same single Service Affect, single Information Control, and single Library as Place items, plus two of the remaining eight (i.e. nine minus the one core item completed by everyone) randomly selected Service Affect items, two of the remaining seven (i.e. eight minus the one core item completed by everyone) randomly selected Information Control, and one of the remaining four (i.e. five minus the one core item completed by everyone) randomly selected Library as Place items.

Here we explore the features of the LibQUAL + [®] Lite protocol implemented at four university libraries in the USA. Specifically, we were interested in exploring participation rates, completion times, and result comparisons across the two protocols at each of the four institutions. An important meta-analysis of the literature conducted by Colleen Cook, Fred Heath, and Russell L. Thompson suggested that response rates should be improved with the use of the shorter protocol (Cook *et al.*, 2000).

In our study, at one institution 70 percent of all LibQUAL + [®] survey invitees were randomly assigned the LibQUAL + [®] Lite protocol, while the remaining 30 percent of

Item	Bob	Mary	Person Bill	Sue	Ted
Service Affect #1	**X**	**X**	**X**	**X**	**X**
Information Control #1	**X**	**X**	**X**	**X**	**X**
Service Affect #2		X		X	
Library as Place #1	**X**	**X**	**X**	**X**	**X**
Service Affect #3	X	X			
Info Control #2				X	X
Library as Place #2			X	X	

Note: Items completed by *all* participants are presented in bold

Table I.
Matrix sampling survey strategy

survey invitees were randomly assigned the long form of the protocol (i.e. all 22 survey core items). At the remaining three institutions, 50 percent of all LibQUAL + ® survey invitees were randomly assigned the LibQUAL + ® Lite protocol, while the remaining 50 percent of survey invitees were randomly assigned the long form of the protocol (i.e. all 22 survey core items).

Results
Survey completion rates
Table II presents the number of participants across the four institutions (assigned ID numbers 433, 3, 107, and 5 to assure their anonymity) and the two protocol forms (i.e. short and long). As indicated in Table II, the actual participation rates for persons randomly assigned the LibQUAL + ® Lite protocol (i.e. 73.7 percent, 59.5 percent, 55.0 percent, and 55.3 percent, respectively) were higher than the baselines of the percentages of persons at each institution asked to complete the short form (i.e. 70 percent, 50 percent, 50 percent, and 50 percent, respectively). Thus, these results clearly indicate that participants are more likely to complete the survey when the matrix sampling strategy is used to collect data on all the items in the item pool.

Table III presents the median survey completion times in seconds across the two protocols. These results indicate that completing the protocol with only eight versus all 22 core items took a little more than half as long as completion of the full item set.

Of course, participants in both groups completed other items (e.g. demographic self-descriptions), and were allowed to provide comments. Historically, about 40 percent of all LibQUAL + ® participants write comments, and these qualitative data are at least as important as the quantitative data gathered on the protocol, because here users often present specific suggestions for library improvement! And persons writing

Group/statistic	Institution			
	433	3	107	5
Short	1,868	627	451	382
Long	688	426	369	309
Total	2,536	1,053	820	691
Actual (percent)	73.7	59.5	55.0	55.3
Random (percent)	70.0	50.0	50.0	50.0
Difference (percent)	3.7	9.5	5.0	5.3

Note: At institution #433, 70 (percent) of all participants were *randomly* assigned the short form, while at the remaining three institutions 50 (percent) of all participants were *randomly* assigned the short form

Table II.
Ratios of completers across the two administration formats and four institutions

Format	Institution				Total
	433	3	107	5	
Long form	456.5	501.0	458.0	470.0	470.5
"Lite" form	276.0	300.0	290.0	291.5	285.0

Table III.
Median completion times in seconds across administration formats

longer comments would have taken longer to complete the survey regardless of which protocol they were randomly assigned.

Table IV presents the percentages of participants who both completed the survey once started and met the protocol inclusion criteria across institutions and administration formats. In LibQUAL + [®], participants are excluded from the dataset if they meet certain criteria. For example, on the longer protocol, if a participant answer more than 11 core items "not applicable", the participant's data are dropped under a view that such a user for whatever reason does not have a definitive view of library service quality. Also, no participant can logically rate a service item higher on what is "minimally acceptable" than the "desired" service quality on the same item, and any person with an excessive number of such "inversions" also is omitted under a view that the person is responding randomly rather than seriously.

The results shown in Table IV make clear that higher percentages of persons who start the LibQUAL + [®] Lite protocol once they begin actually complete the survey. Of course, some participants begin the survey, determine what the protocol is about, and return later to actually complete the survey. So, not all persons who fail to complete the survey in a given administration are actually non-responders.

LibQUAL + [®] Lite versus LibQUAL + [®] score comparisons

The present study was a randomized clinical trial or an experiment. Because the participants were randomly assigned either the LibQUAL + [®] Lite protocol or the conventional LibQUAL + [®] protocol, the scores on the measures at a given institution should be similar, unless the protocols themselves caused score differences.

All LibQUAL + [®] scores (i.e. total, the three scales, and items) are scaled on a scale of 1-9, with 9 being the highest rating. Figure 1 presents 95 percent confidence intervals

Table IV.
Percentages of participants who both completed the survey once started and met inclusion criteria across institutions and administration formats

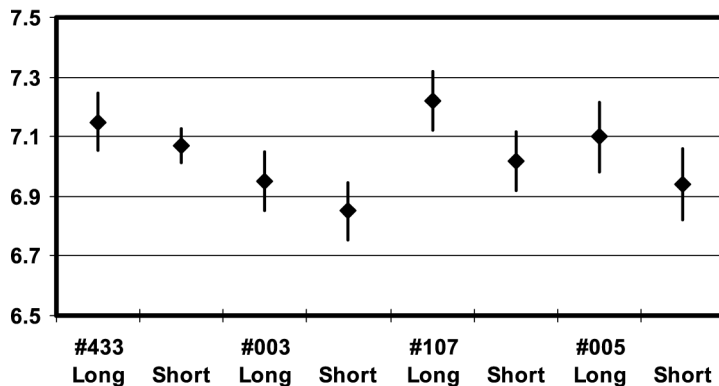|	Institution				
Format	433	3	107	5	Total
Long form	56.18	35.26	61.40	51.07	49.18
"Lite" form	66.08	51.44	73.57	60.54	62.91

Figure 1.
Ninety-five percent confidence intervals about means for total perception scores across four institutions and two protocols

around the means of the total perception scores across the four institutions and the two protocols. Total perception score means across institutions should not necessarily be equal, but means across protocols *within* institutions would be expected to be equal, or at least similar, unless the protocol variation caused score differences.

The results shown in Figure 1 show that the mean total perception scores were somewhat lower on the LibQUAL $+$ [®] Lite protocol than on the conventional protocol at all four institutions. The greatest difference occurred at institution #107. Thus, the samples of people deciding to provide data are to some extent qualitatively different across the two protocols, with the samples for the Lite protocol having somewhat more negative views of library service quality.

Figure 2 presents 95 percent confidence intervals around the means of the perception scores on the Service Affect (top graph in Figure 2), Information Control, and Library as Place (bottom graph in Figure 2). Again, means on the LibQUAL $+$ [®] Lite protocol tended to be somewhat lower than those on the conventional protocol, although the means on the Service Affect scale were reasonably similar. Differences were most pronounced on the Library as Place scale and at institution #107.

Figure 3 presents 95 percent confidence intervals around the means of the item scores on the three items (i.e. one from each of the three LibQUAL $+$ [®] scales) completed by *all* participants, including those randomly assigned the Lite protocol. Again, means tended to be similar or somewhat lower on the Lite protocol, except on the Service Affect item, and the greatest mean differences occurred at institution #107.

Discussion

Our results show that higher completion rates are associated with the shorter survey completion times that can be achieved by randomly selecting which items will be completed by given participants. At the same time, data can be collected on all survey items, even though demands on individual participants are minimized by the LibQUAL $+$ [®] Lite protocol.

The somewhat troubling aspects of our results is that the samples realized by the two protocols may be qualitatively different. Theoretically, because the protocols were randomly assigned, the total, scale, and item means should be comparable across the two protocols at a given institution, unless the Lite protocol itself yields participants with different views of library service quality.

Implications

Because greater completion rates were realized with the LibQUAL $+$ [®] Lite protocol, the data from the Lite protocol might be the most accurate representation of the views of all the library users in a given community. Thus, our results might be interpreted as meaning that when the longer protocol is used, persons with more positive views are disproportionately likely to respond to the survey. This means that the longer protocol might result in somewhat inflated ratings of library service quality.

If this is true, our results argue for the adoption of the LibQUAL $+$ [®] Lite protocol as against the traditional LibQUAL $+$ [®] protocol. Thus, in future years we may routinely encourage libraries to present only the LibQUAL $+$ [®] Lite protocol to their users, or at least to a large randomly selected percentage of their users. This approach would also have the benefit of mitigating time demands on users, while still allowing collection of data on all core items.
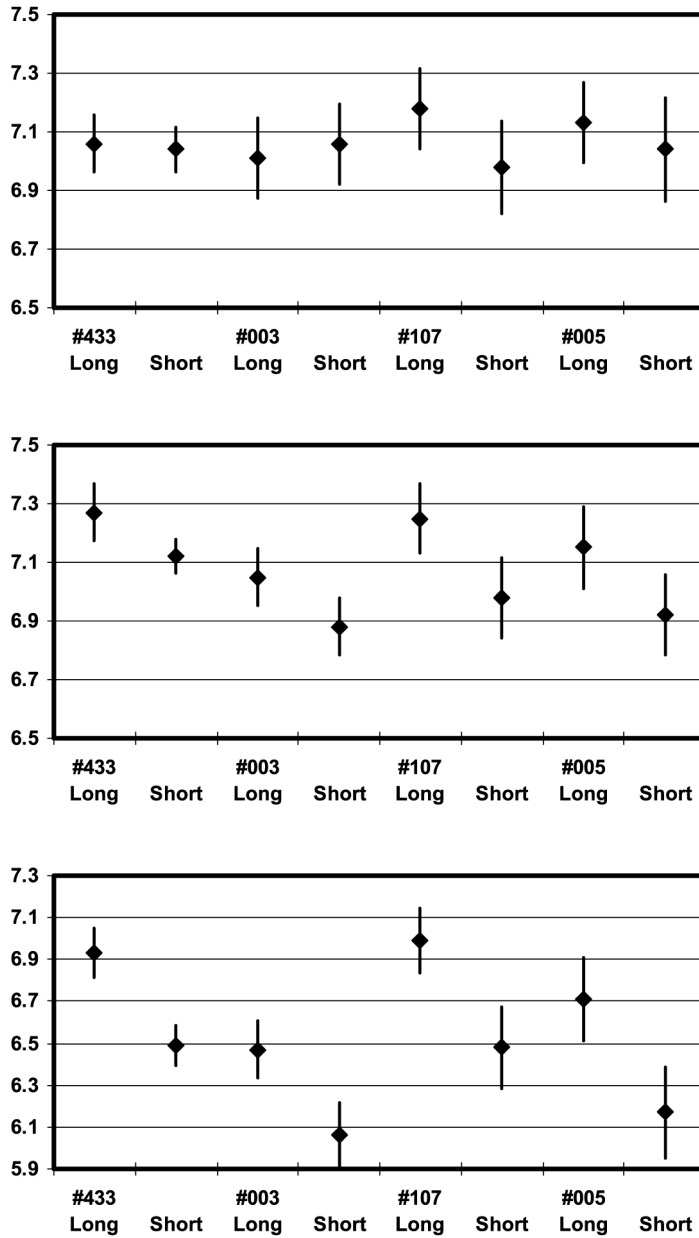
Figure 2.
Ninety-five percent
confidence intervals about
means for Service Affect
(top), Information Control
(middle), and Library as
Place (bottom) scale scores
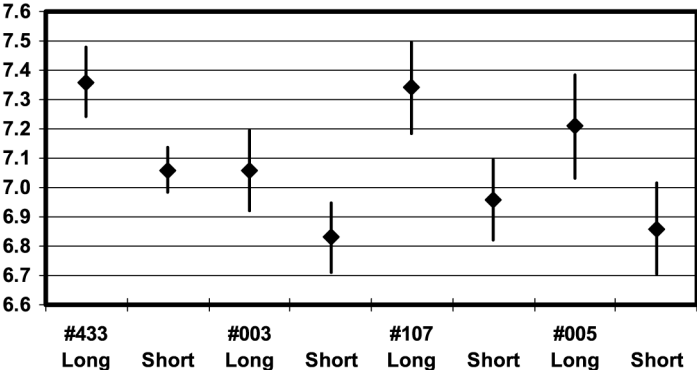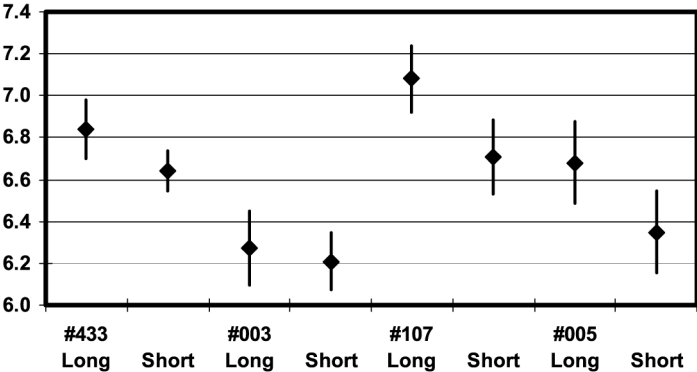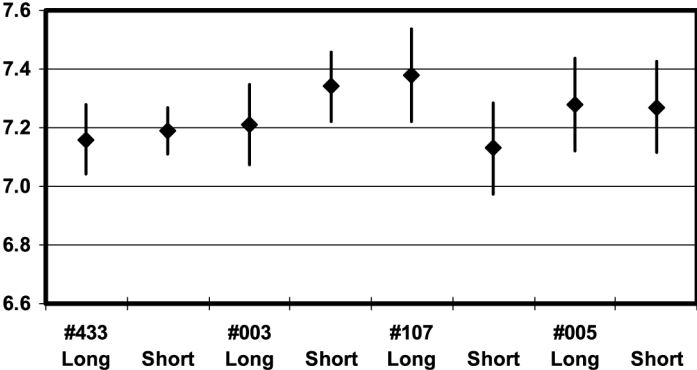across four institutions
and two protocols

Figure 3.
Ninety-five percent
confidence intervals about
means for Service Affect
(top), Information Control
(middle), and Library as
Place (bottom) scores on
items 13, 10 and 3 across
four institutions and two
protocols

Equating scores across the two protocols
The apparent finding that the LibQUAL + ® Lite protocol and the traditional
LibQUAL + ® protocol yield somewhat different ratings does *not* mean that scores on
the two protocols cannot be equated either within a given institution at which some
participants complete the long form and others the Lite form, or across years in which
the long form is used one year and the Lite form is used in another year. The three
"linking" items completed by all participants in both formats provide the necessary
linking equations.

The equation to convert scores on the Service Affect items on the traditional
LibQUAL + ® protocol into the metric of the LibQUAL + ® Lite protocol is:

$$(\text{Service Affect item scores} + 0.03) \times 1.01. \tag{1}$$

Conversely, the equation to convert scores on the Service Affect items on the
LibQUAL + ® Lite protocol into the metric of the traditional LibQUAL + ® protocol is:

$$(\text{Service Affect item scores} - 0.03) \times 0.99. \tag{2}$$

The equation to convert scores on the Information Control items on the traditional
LibQUAL + ® protocol into the metric of the LibQUAL + ® Lite protocol is:

$$(\text{Information Control item scores} - 0.27) \times 1.01. \tag{3}$$

Conversely, the equation to convert scores on the Information Control items on the
LibQUAL + ® Lite protocol into the metric of the traditional LibQUAL + ® protocol is:

$$(\text{Information Control item scores} + 0.27) \times 0.99. \tag{4}$$

The equation to convert scores on the Library as Place items on the traditional
LibQUAL + ® protocol into the metric of the LibQUAL + ® Lite protocol is:

$$(\text{Library as Place item scores} - 0.32) \times 1.10. \tag{5}$$

Conversely, the equation to convert scores on the Library as Place items on the
LibQUAL + ® Lite protocol into the metric of the traditional LibQUAL + ® protocol is:

$$(\text{Library as Place item scores} + 0.32) \times 0.95. \tag{6}$$

These equations allow libraries administering the Lite protocol to compare scores
apples-to-apples contemporaneously with libraries *not* using the Lite form, or *vice
versa*. The equations also allow a given library now administering the Lite protocol to
compare current scores apples-to-apples retrospectively with the same institution's
prior scores obtaining using the traditional LibQUAL + ® protocol.

Conclusion
One reason for the success of the LibQUAL + ® project is that the four developers –
Fred Heath, Colleen Cook, Martha Kyrillidou, and Bruce Thompson – have
meticulously studied and documented myriad features of the protocol. For example,
issues such as the use of radio buttons versus slider bars on the web-administered
protocol (Cook *et al.*, 2001b) and whether or not users think about library service
quality continuously as matters of degree versus categorically (i.e. there are good

libraries and bad libraries) (Arnau *et al.*, 2001), have been investigated. The present study adds to this literature.

Our results demonstrate that matrix sampling strategies for administering randomly selected items from a larger item pool to a given survey participant has numerous advantages (e.g. fewer time demands on survey participants, higher survey completion rates). These results also demonstrate the value of using "linking" items common to both protocol formats so that formulas can be developed to equate scores across the administration protocols.

Our findings have implications beyond the LibQUAL + ® protocol. The matrix sampling procedure can be usefully applied with a variety of service quality assessments, including locally developed surveys involving numerous assessment items. However, our results also suggest that caution must be exercised with respect to assumptions that protocol changes will not affect the tenor of the assessment data that the protocols yield.

References

Arnau, R.C., Thompson, R.L. and Cook, C. (2001), "Do different response formats change the latent structure of responses? An empirical investigation using taxonometric analysis", *Educational and Psychological Measurement*, Vol. 61, pp. 23-44.

Cook, C. and Heath, F. (2001), "Users' perceptions of library service quality: a 'LibQUAL + ™' qualitative study", *Library Trends*, Vol. 49, pp. 548-84.

Cook, C., Heath, F. and Thompson, B. (2001a), "Users' hierarchical perspectives on library service quality: a 'LibQUAL + ™' study", *College and Research Libraries*, Vol. 62, pp. 147-53.

Cook, C., Heath, F. and Thompson, B. (2002), "Score norms for improving library service quality: a LibQUAL + ™ study", *Portal: Libraries and the Academy*, Vol. 2, pp. 13-26.

Cook, C., Heath, F. and Thompson, B. (2003), "'Zones of tolerance' in perceptions of library service quality: a LibQUAL + ™ study", *portal: Libraries and the Academy*, Vol. 3, pp. 113-23.

Cook, C., Heath, F. and Thompson, R.L. (2000), "A meta-analysis of response rates in web- or internet-based surveys", *Educational and Psychological Measurement*, Vol. 61, pp. 821-36.

Cook, C., Heath, F., Thompson, R.L. and Thompson, B. (2001b), "Score reliability in web- or internet-based surveys: unnumbered graphic rating scales versus Likert-type scales", *Educational and Psychological Measurement*, Vol. 61, pp. 697-706.

Cook, C., Heath, F., Kyrillidou, M., Thompson, B. and Thompson, R.L. (2001c), "The search for new measures: the ARL LibQUAL + ™ study – a preliminary report", *portal: Libraries and the Academy*, Vol. 1, pp. 103-12.

Cullen, R.J. (2001), "Perspectives on user satisfaction surveys", *Library Trends*, Vol. 49 No. 4, pp. 662-86.

de La Rochefoucauld, F. (1613-1680), "Maxim 550 (G.E.F.)", *Maximes* ("It is more necessary to study men than books").

Heath, F., Cook, C., Kyrillidou, M. and Thompson, B. (2002), "ARL Index and other validity correlates of LibQUAL + ™ scores", *portal: Libraries and the Academy*, Vol. 2, pp. 27-42.

Nitecki, D.A. (1996), "Changing the concept and measure of service quality in academic libraries", *The Journal of Academic Librarianship*, Vol. 22 No. 3, pp. 181-90.

Thompson, B. (2006), "Measuring user perceptions of library service quality: an introduction to LibQUAL + ™", paper presented at Czech and Slovakian Library Information Network (CASLIN) Conference, Prague, June.

Thompson, B. (2007), "The origins/birth of LibQUAL + ®", available at: www.coe.tamu.edu/~bthompson/libbirth.htm

Thompson, B., Cook, C. and Heath, F. (2001), "How many dimensions does it take to measure users' perceptions of libraries? A LibQUAL + ™ study", *portal: Libraries and the Academy*, Vol. 1, pp. 129-38.

Thompson, B., Cook, C. and Kyrillidou, M. (2005), "Concurrent validity of LibQUAL + ™ scores: what do LibQUAL + ™ scores measure?", *Journal of Academic Librarianship*, Vol. 31, pp. 517-22.

Thompson, B., Kyrillidou, M. and Cook, C. (2007), "On-premises library versus Google™-like information gateway usage patterns: a LibQUAL + ™ study", *portal: Libraries and the Academy*, Vol. 7, pp. 463-80.

Thompson, B., Kyrillidou, M. and Cook, C. (2008), "Library users' service desires: a LibQUAL + ™ study", *Library Quarterly*, Vol. 1, pp. 1-18.

Zeithaml, V.A., Parasuraman, A. and Berry, L.L. (1990), *Delivering Quality Service: Balancing Customer Perceptions and Expectations*, The Free Press, New York, NY.

Further reading

Association of Research Libraries (n.d.), "DigiQUAL + ®", available at: www.digiqual.org/digiqual/index.cfm

Corresponding author

Bruce Thompson can be contacted at: bruce-thompson@tamu.edu