Scaling Users' Perceptions of Library Service Quality Using Item Response Theory: A LibQUAL+™ Study

Youhua Wei, Bruce Thompson, and C. Colleen Cook

abstract: LibQUAL+™ data to date have not been subjected to the modern measurement theory called polytomous item response theory (IRT). The data interpreted here were collected from 42,090 participants who completed the "American English" version of the 22 core LibQUAL+™ items, and 12,552 participants from Australia and Europe who completed the "British English" version of the 22 core LibQUAL+™ items. Results suggest that the LibQUAL+™ protocol has psychometric integrity, that American English participants tend to be somewhat less critical of library service quality, and that students and faculty have fairly similar views of library service quality.

As Amos Lakos and Shelley Phipps noted in their recent *portal: Libraries and the Academy* article, "Libraries must develop internal organizational systems that enable successful assessment and evaluation of their services and processes to achieve positive outcomes for customers. . . . A key implication of this focus is the realization that *what gets measured gets managed* [italics added]; creating value lies in understanding, communicating, and measuring what matters to customers."[1] Clearly, in the Internet era, collection counts are no longer a sufficient index of library quality.

> **Clearly, in the Internet era, collection counts are no longer a sufficient index of library quality.**

The LibQUAL+™ protocol was developed as one of the "New Measures" initiatives of the Association of Research Libraries and its various partners. More than 100,000 library users from the United States, Europe, and Australia have now completed the spring 2004 LibQUAL+™ survey.

The Web-administered survey consists, in part, of 22 core items measuring three dimensions of library service: affect of service, library as place, and information control. Additionally, participants have the option of providing open-ended remarks in a comments box, and historically roughly 40 percent of users do so. Thus, today researchers refer to LibQUAL+™ as "22 items plus a box."

Designed to measure user satisfaction with library service quality, the LibQUAL+™ program is now in its fifth year of operation. Over the past five years, LibQUAL+™ has been tested in every state but two, as well as in Canada, Australia, Egypt, England, France, Ireland, Scotland, Sweden, the Netherlands, and the United Arab Emirates. Data have been collected from over 300,000 users. The current survey instrument is available in eight language variations.

The growing LibQUAL+™ community of participants and its robust dataset provide rich resources for analyzing and improving library services. LibQUAL+™ results have been widely published in the scholarly journals dealing with librarianship (e.g., *College and Research Libraries*, *IFLA Journal*, *Journal of Library Administration*, *Library Administration & Management*, *Library Quarterly*, *Library Trends*, *Performance Measurement and Metrics*, and *portal: Libraries and the Academy*).[2]

Various quantitative methods have been brought to bear on LibQUAL+™ data to confirm that scores have reasonable psychometric integrity. As explained by Bruce Thompson, it is essential to determine (a) whether the scores measure something (versus being random), which is the question of score *reliability*, and (b) whether the extent to which the scores measure something, this something is the correct something, which is the question of score *validity*.[3] An impressive array of LibQUAL+™ reliability and validity results have been published.[4]

But extremely complex and sophisticated statistical methods have also been used to adduce evidence regarding the psychometric integrity of LibQUAL+™ scores. Among these methods have been reliability generalization and taxonometric analyses.[5]

However, LibQUAL+™ data to date have not been subjected to the modern measurement theory called polytomous item response theory (IRT). IRT methods are extremely complicated but have been explained in accessible ways by several authors.[6] As explained in more detail elsewhere, among other purposes IRT analyses seek to quantify the degree to which various items function well across a range of respondents with different views (called item discrimination) and which items yield the most information for respondents holding a given view of libraries (e.g., highly unfavorable, moderately favorable) along the continuum of possible ratings of library service quality (called location).

IRT methods were employed here to address the second of our two research questions:

1. Do perceptions of library service quality differ across user groups or across the libraries that completed the "American English" versus the "British English" versions of LibQUAL+™?
2. Across the "American English" versus the "British English" version of LibQUAL+™, how do the 22 core quantitative items scale as measures of perceived library service quality?

IRT was particularly relevant to investigating the second question, because theoretically the methods yield the same discrimination and location statistics across different

groups of respondents (i.e., these IRT statistics are theoretically "person free"). If the statistics varied dramatically across user groups (e.g., American versus English), such variations would suggest measurement problems with the protocol. Only IRT analyses have these theoretical features.

Participants

The data interpreted here were collected during spring 2004 from undergraduate and graduate students and faculty. Data were collected via the Web from 42,090 participants who completed the original "American English" version of the 22 core LibQUAL+™ items. There were more females (59.0 percent) than males in the sample. These 42,090 participants came from 156 institutions.

Data were also collected from 12,552 participants who completed the "British English" version of the 22 core LibQUAL+™ items. Thirteen of the 22 items had the same wording across these two language versions. Library staff at participating institutions suggested wording changes for the remaining nine items. There were more females (59.9 percent) than males in the sample. These 12,552 participants came from 18 institutions in Australia and Europe who elected to use this version of the measure, rather than one of the other language translations (e.g., Dutch, Swedish, Continental French) not considered here.

The gender breakdowns describe the makeup of our sample for comparative purposes when new data are collected and also suggest that these two samples were relatively comparable regarding gender representation. Table 1 presents the numbers of students and faculty in the three role groups across the two samples.

Results

Research Question #1

Table 2 presents the mean perception ratings across the three role groups (i.e., undergraduate students, graduate students, and faculty) and the two language groups (i.e., American English and British English). On the LibQUAL+™ protocol, 1 is the lowest possible rating, and 9 is the highest possible rating. The tabled entries are sorted from highest mean American English rating (i.e., "employees are consistently courteous") to lowest rating (i.e., "employees instill confidence in users").

Table 2 also presents the standard deviations (SDs) of the mean ratings of the three role groups. If the three means were identical, this SD would be zero. For example, in both language samples the three role groups had virtually the same ratings on the item, "easy access tools that allow me to find things on my own." The SD gets larger as the means increasingly diverge.

Research Question #2

The polytomous IRT analyses were conducted using the PARSCALE 4 software.[7] Table 3 presents the IRT coefficients for the 22 core LibQUAL+™ items for the American English (n_1 = 42,090) sample. Table 4 presents the IRT coefficients for the British English (n_2 = 12,552) sample.

Table 1

Participant Breakdown by Role Groups

Role Group	American Eng.		British Eng.	
	n	%	*n*	%
Undergraduate	24,408	58.0	8,668	69.1
Graduate Student	10,727	25.5	2,793	22.3
Faculty	6,955	16.5	1,091	8.7
Total	42,090		12,552	

IRT coefficients are always presented as logarithmic values called "logits." These coefficients are especially useful in statistical analyses, because in theory they best generalize to new samples.

One coefficient is called "location." Location measures where an item is scaled regarding overall perceptions of library service quality. Larger values indicate items on which persons who have higher overall perception scores are more likely to rate a given item higher (i.e., more favorably).

The second coefficient is called "discrimination." This measures the extent to which an item does well at measuring opinions across a range of persons holding different attitudes about library service quality.

Discussion

Our two research questions had different purposes. The first question had the purpose of exploring differences in perceptions of North American versus other library users and across student and faculty user groups in order to isolate items that users tend to have similar opinions on or diverge in their perceptions. The focus of the second research question was not descriptive but, instead, focused on the psychometric integrity of LibQUAL+™.

First, it is striking that across the three respondent groups the American English ratings were higher than the British English ratings on all 22 LibQUAL+™ core items.

Research Question #1

The study's first research question asked, "Do perceptions of library service quality differ across user groups or across the libraries that completed the 'American English' versus the 'British English' versions of LibQUAL+™?" The results presented in table 2 suggest a number of conclusions.

First, it is striking that across the three respondent groups the American English ratings were higher than the British English ratings on all 22 LibQUAL+™ core items.

Table 2

Means (and Standard Deviations) of Service Quality Perceptions Across Role Groups and Language Samples

Item	Item Content Core	Role Group				Rank	SD	Inter Group
		Undergrad.	Grad Stud.	Faculty	Total			
AS06	**Employees are consistently courteous**	7.12(1.75)	7.30(1.66)	7.67(1.50)	7.26(1.68)	**1**		**0.23**
	Library staff who are consistently courteous	6.65(1.81)	6.87(1.78)	7.39(1.56)	6.77(1.78)		4	0.31
IC14	*Modern equipment lets me easily access needed info.*	7.27(1.56)	7.21(1.52)	7.15(1.49)	7.24(1.54)	**2**		**0.05**
		6.70(1.59)	6.68(1.61)	6.72(1.45)	6.69(1.59)		5	0.02
AS18	*Willingness to help users*	7.13(1.63)	7.26(1.55)	7.54(1.50)	7.23(1.59)	**3**		**0.17**
		6.62(1.61)	6.77(1.66)	7.06(1.57)	6.69(1.62)		6	0.18
AS09	**Readiness to respond to users' questions**	7.10(1.62)	7.27(1.54)	7.55(1.50)	7.22(1.58)	**4**		**0.18**
	Readiness to respond to users' enquiries	6.73(1.61)	6.80(1.68)	7.21(1.53)	6.79(1.62)		3	0.21
AS11	**Employees have knowledge to answer user questions**	7.14(1.62)	7.15(1.59)	7.35(1.56)	7.18(1.60)	**5**		**0.09**
	Staff have knowledge to answer user questions	6.87(1.57)	6.88(1.58)	7.03(1.61)	6.88(1.58)		1	0.07
IC19	*Making info easily access for independent use*	7.13(1.52)	7.16(1.47)	7.19(1.47)	7.15(1.50)	**6**		**0.02**
		6.62(1.51)	6.70(1.54)	6.86(1.40)	6.66(1.51)		8	0.10
IC05	*Library Web site enabling locate info. on my own*	7.12(1.67)	7.14(1.60)	7.07(1.63)	7.12(1.65)	**7**		**0.03**
		6.77(1.67)	6.83(1.64)	6.99(1.54)	6.80(1.65)		2	0.09
AS13	**Employees deal with users in a caring fashion**	7.05(1.64)	7.08(1.61)	7.36(1.55)	7.11(1.62)	**8**		**0.14**
	Staff who deal with users in a caring fashion	6.41(1.69)	6.58(1.70)	6.82(1.61)	6.48(1.69)		12	0.17

Table 2 cont.

Item	Item Content Core	Role Group				Rank	SD	Inter Group
		Undergrad.	Grad Stud.	Faculty	Total			
IC16	*Easy access tools allow find things on my own*	**7.05(1.57)**	**7.09(1.49)**	**7.06(1.50)**	**7.06(1.54)**	**9**	9	**0.02**
		6.65(1.55)	6.70(1.56)	6.79(1.46)	6.67(1.54)		7	0.06
AS15	**Employees understand the needs of their users**	**7.01(1.61)**	**7.04(1.58)**	**7.19(1.61)**	**7.04(1.60)**	**10**		**0.08**
	Staff who understand the needs of their users	6.54(1.59)	6.60(1.66)	6.93(1.64)	6.59(1.61)		9	0.17
IC10	*The electronic information resources I need*	**7.03(1.55)**	**7.00(1.55)**	**7.01(1.59)**	**7.02(1.56)**	**11**		**0.02**
		6.43(1.55)	6.54(1.60)	6.80(1.47)	6.49(1.55)		11	0.15
AS22	*Dependability handling users' service problems*	**6.95(1.64)**	**7.01(1.61)**	**7.23(1.57)**	**7.01(1.62)**	**12**		**0.12**
		6.49(1.56)	6.55(1.63)	6.86(1.57)	6.54(1.58)		10	0.16
LP12	*A comfortable and inviting location*	**7.11(1.74)**	**6.77(1.84)**	**6.78(1.84)**	**6.97(1.78)**	**13**		**0.16**
		6.16(1.87)	6.10(1.85)	5.75(1.83)	6.11(1.83)		18	0.18
IC02	*Making elec. resources access home or office*	**6.90(1.81)**	**6.96(1.79)**	**7.97(1.79)**	**6.92(1.82)**	**14**		**0.49**
		6.34(1.89)	6.55(1.85)	6.71(1.74)	6.42(1.87)		13	0.15
LP17	**A getaway for study, learning, or research**	**7.03(1.70)**	**6.71(1.74)**	**6.53(1.82)**	**6.86(1.73)**	**15**		**0.21**
	A haven for study, learning, or research	6.14(1.71)	5.96(1.80)	5.56(1.86)	6.05(1.74)		19	0.24
IC20	*Print and/or elec. journal coll. require for work*	**6.94(1.70)**	**6.83(1.73)**	**6.59(1.83)**	**6.85(1.73)**	**16**		**0.14**
		6.30(1.73)	6.32(1.81)	6.43(1.70)	6.32(1.74)		14	0.06
LP08	**Quiet space for individual activities**	**6.95(1.83)**	**6.67(1.87)**	**6.58(1.84)**	**6.82(1.84)**	**17**		**0.16**
	Quiet space for individual work	6.25(1.93)	5.89(2.00)	5.57(1.90)	6.11(1.94)		17	0.28
IC07	*The printed library materials I need for work*	**6.85(1.69)**	**6.74(1.68)**	**6.51(1.81)**	**6.77(1.71)**	**18**		**0.14**
		6.14(1.67)	6.15(1.75)	6.29(1.67)	6.16(1.69)		16	0.07

Item	Description							
LP03	*Library space that inspires study and learning*	6.78(1.85)	6.38(1.96)	6.39(1.93)	6.61(1.89)	**19**		**0.19**
		5.85(1.86)	5.68(1.92)	5.55(1.85)	5.78(1.87)		21	0.12
LP21	**Community space for group learning and group study**	6.77(1.84)	6.43(1.84)	6.07(1.89)	6.57(1.85)	**20**		**0.29**
	Space for group learning and group study	5.53(2.12)	5.45(2.09)	4.76(2.07)	5.44(2.11)		22	0.34
AS04	*Giving users individual attention*	6.37(1.86)	6.65(1.75)	7.11(1.65)	6.56(1.80)	**21**		**0.30**
		5.66(1.78)	6.00(1.79)	6.39(1.71)	5.80(1.78)		20	0.30
AS01	**Employees instill confidence in users**	6.35(1.72)	6.54(1.72)	6.96(1.66)	6.50(1.71)	**22**		**0.25**
	Library staff who instill confidence in users	6.04(1.63)	6.31(1.63)	6.74(1.58)	6.16(1.62)		15	0.29

Note. Items presented in *italics* were worded identically for both the "American English" and the "British English" survey versions. Results presented in **bold** were for the American English sample. The last two digits of the Item code present the sequencing of the 22 LibQUAL+™ items within the survey. The first two characters of the Item code indicate which scale a given item measures (i.e., "AS" = Affect of Service; "IC" = Information Control; "LP" = Library as Place). Standard Deviations for a given group means are presented in parentheses.

Table 3

IRT Coefficients for the **American English** ($n_1 = 42{,}090$) Sample

Item	Item Content Core	Coefficient	
		Discrimination	**Location**
AS06	Employees are consistently courteous	1.197	-1.515
IC14	*Modern equipment lets me easily access needed info.*	1.377	-1.452
AS09	Readiness to respond to users' questions	1.495	-1.443
AS18	*Willingness to help users*	1.594	-1.440
AS11	Employees have knowledge to answer user questions	1.501	-1.408
IC05	*Library Web site enabling locate info. on my own*	1.143	-1.386
IC19	*Making info easily access for independent use*	1.673	-1.361
AS13	Employees deal with users in a caring fashion	1.548	-1.354
IC16	*Easy access tools allow find things on my own*	1.524	-1.302
LP12	*A comfortable and inviting location*	1.067	-1.298
AS15	Employees understand the needs of their users	1.688	-1.291
IC02	*Making elec. resources access home or office*	0.963	-1.284
AS22	*Dependability handling users' service problems*	1.515	-1.276
IC10	*The electronic information resources I need*	1.408	-1.276
LP17	A getaway for study, learning, or research	1.223	-1.208
IC20	*Print and/or elec. journal coll. require for work*	1.188	-1.188
LP08	Quiet space for individual activities	0.987	-1.184
IC07	*The printed library materials I need for work*	1.215	-1.121
LP03	*Library space that inspires study and learning*	0.984	-1.081
LP21	Community space for group learning and group study	1.016	-1.014
AS04	*Giving users individual attention*	1.189	-0.978
AS01	Employees instill confidence in users	1.259	-0.899

Note. Items presented in *italics* were worded identically for both the "American English" and the "British English" survey versions. The last two digits of the Item code present the sequencing of the 22 LibQUAL+™ items within the survey. The first two characters of the Item code indicate which scale a given item measures (i.e., "AS" = Affect of Service; "IC" = Information Control; "LP" = Library as Place).

The basis for this result is not entirely clear. The result might occur if libraries in the British English sample had fewer resources or if users in these libraries were more demanding.

However, the results for item 2, an information control item, may shed some light on this issue. The greatest difference (intergroup SD = 0.42) in the mean ratings across the three role groups occurred in the American English sample for the item, "making electronic resources accessible from home or office." The ratings in this sample were

Table 4

IRT Coefficients for the **British English** (n_2 = 12,552) Sample

Item	Item Content Core	Coefficient Discrimination	Location
AS11	Staff have knowledge to answer user questions	1.181	-1.408
IC05	*Library Web site enabling locate info. on my own*	0.918	-1.369
AS06	Library staff who are consistently courteous	0.925	-1.360
AS09	Readiness to respond to users' enquiries	1.166	-1.324
IC14	*Modern equipment lets me easily access needed info.*	1.139	-1.240
AS18	*Willingness to help users*	1.302	-1.235
IC16	*Easy access tools allow find things on my own*	1.213	-1.210
IC19	*Making info easily access for independent use*	1.372	-1.188
AS15	Staff who understand the needs of their users	1.307	-1.141
AS22	*Dependability handling users' service problems*	1.261	-1.098
IC02	*Making elec. resources access home or office*	0.759	-1.077
AS13	Staff who deal with users in a caring fashion	1.198	-1.063
IC10	*The electronic information resources I need*	1.150	-1.057
IC20	*Print and/or elec. journal coll. require for work*	1.001	-0.950
LP08	Quiet space for individual work	0.780	-0.830
IC07	*The printed library materials I need for work*	1.049	-0.809
LP12	*A comfortable and inviting location*	0.947	-0.790
AS01	Library staff who instill confidence in users	1.093	-0.788
LP17	A haven for study, learning, or research	1.078	-0.740
LP03	*Library space that inspires study and learning*	0.867	-0.527
AS04	*Giving users individual attention*	0.996	-0.524
LP21	Space for group learning and group study	0.727	-0.314

Note. Items presented in *italics* were worded identically for both the "American English" and the "British English" survey versions. The last two digits of the Item code present the sequencing of the 22 LibQUAL+™ items within the survey. The first two characters of the Item code indicate which scale a given item measures (i.e., "AS" = Affect of Service; "IC" = Information Control; "LP" = Library as Place).

considerably higher in the faculty (M_F = 7.97) than in the undergraduate and graduate student groups (M_U = 6.90; M_G = 6.96).

The ratings in the British English sample were both lower and more homogenous than the ratings in the two student groups (M_U = 6.34; M_G = 6.55; M_F = 6.71). These patterns would make sense if (a) American libraries do have more resources, and (b) American faculty are more aware than students about the availability of on-line services that these greater resources have produced.

Second, the rank order of the 22 item means was reasonably similar across the two language samples, as reported in table 2. Thus, although the British English participants were somewhat more critical in their views, participants in both samples tended to agree about what was perceived as going better or worse within their libraries.

Research Question #2

The study's second research question asked, "Across the 'American English' versus the 'British English' version of LibQUAL+™, how do the 22 core quantitative items scale as measures of perceived library service quality?" As reported in tables 3 and 4, the location coefficients for the 22 core items indicate that the items scaled in roughly the same order as the means reported in table 2.

> **LibQUAL+™ users need not be overly concerned that their results are some artifact of using means from a measurement model that is too simplistic.**

That is, even a sophisticated measurement theory, such as item response theory, does not yield scaling results that differ from a simple inspection of the means. Thus, using means to report and interpret LibQUAL+™ results appears sensible. LibQUAL+™ users need not be overly concerned that their results are some artifact of using means from a measurement model that is too simplistic.

The discrimination coefficients also present encouraging information regarding score quality. The coefficients were reasonably homogeneous, indicating that all 22 items do approximately as well at yielding scores across people with differing views of the service quality at their libraries.

Some Final Thoughts

In their recent *portal* article, Lakos and Phipps noted that library "assessment activities also require certain skills more aligned to marketing and business than to librarianship. Assessment has not been taught or appreciated by the profession."[8] But the profession is changing.

The impact of the LibQUAL+™ project has been considerable. And it is important that evidence of score integrity has regularly been presented by project workers. But ultimately the project's impact on creating a culture of assessment within libraries will occur—not as a function of the protocol itself, but as a by-product of the assessment-related modeling and training required by the project. These impacts will last beyond the lifetime of the protocol itself.

Administration of LibQUAL+™ during the past five years has required that ARL and Texas A&M University staff conduct assessment-related training for hundreds of librarians both before and after administration of the survey. The weeklong, intensive, annual Service Quality Assessment Academy, to be offered for the fourth time in 2005, has also been an important venue for developing new skills sets for librarians.[9]

Equally important have been the initiatives of librarians working at various institutions to explore ways to translate LibQUAL+™ assessment data into action, described

by them in various articles.[10] An especially exciting mechanism for executing this translation is the Summit, initially developed at Clemson University and used several times there and also used at the University of Texas.[11]

Assessment protocols, such as LibQUAL+™, sometimes raise more questions than they answer. Although roughly 40 percent of participants provide open-ended comments explaining their views and making policy recommendations, sometimes focus groups must be conducted to elaborate policy implications.

The Summit is essentially a large focus group, involving various user groups (undergraduate and graduate students and faculty) but also policy-makers, such as university presidents, provosts, and deans.[12] So, the Summit protocol, in addition to *(a) communicating LibQUAL+TM results and fleshing out assessment data, is at once (b) also a vehicle to bring policy-makers together to formulate actions to address assessment data.* The cited ARL Web pages provide more details on the Summit strategy and describe consultation available to assist directors in executing this innovation.

Youhua Wei is a doctoral candidate, University of Florida, Gainesville, FL; he may be contacted via e-mail at: wei2004@ufl.edu.

Bruce Thompson is professor of educational psychology and distinguished research scholar, Texas A & M University, College Station, TX and adjunct professor of family and community medicine, Baylor College of Medicine, Houston, TX; he may be contacted via e-mail at brucethompson@tamu.edu.

C. Colleen Cook is dean of libraries and holder of the Sterling C. Evans Chair, Texas A & M University, College Station, TX; she may be contacted via e-mail at: ccook@tamu.edu.

Notes

1. Amos Lakos and Shelley Phipps, "Creating a Culture of Assessment: A Catalyst for Organizational Change," *portal: Libraries and the Academy* 4, 3 (July 2004): 346–8.
2. Colleen Cook, Fred Heath, and Bruce Thompson, "Users' Hierarchical Perspectives on Library Service Quality: A 'LibQUAL+™' Study," *College and Research Libraries* 62, 2 (March 2001): 147–53; Colleen Cook, Fred Heath, Bruce Thompson, and Russell L. Thompson, "LibQUAL+™: Service Quality Assessment in Research Libraries," *IFLA Journal* 27, 4 (2001): 264–8; Fred Heath, Martha Kyrillidou, and Consuella Askew, eds., "Libraries Act on Their LibQUAL+™ Findings: From Data to Action," special issue, *Journal of Library Administration* 40, 3/4 (Fall 2004); C. A. Snyder, "Measuring Library Service Quality with a Focus on the LibQUAL+™ Project: An Interview with Fred Heath," *Library Administration & Management* 16 (Winter 2002): 4–7; Bruce Thompson, Colleen Cook, and Fred Heath, "Two Short Forms of the LibQUAL+™ Survey Assessing Users' Perceptions of Library Service Quality," *Library Quarterly* 73 (April 2003): 453–65; Colleen Cook and Fred Heath, "Users' Perceptions of Library Service Quality: A 'LibQUAL+™' Qualitative Study," *Library Trends* 49, 4 (Spring 2001): 548–84; Colleen Cook, guest ed., "The Maturation of Assessment in Academic Libraries: The Role of LibQUAL+™," special issue, *Performance Measurement and Metrics* 3, 2 (May 2002); 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, 1 (January 2001): 65–74; 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, 2 (April 2001): 129-38; Colleen Cook, Fred Heath, and Bruce Thompson, "'Zones of Tolerance' in Perceptions of Library Service Quality: A LibQUAL+™ Study," *portal: Libraries and the Academy* 3, 1 (January 2003): 113–23; Fred Heath, Colleen Cook, Martha Kyrillidou, and Bruce Thompson, "ARL Index and Other Validity Correlates of LibQUAL+™ Scores," *portal: Libraries and the Academy* 2, 1 (January 2002): 27–42; and Colleen Cook, Fred Heath, and Bruce Thompson, "Score Norms for Improving Library Service Quality: A LibQUAL+™ Study," *portal: Libraries and the Academy* 2, 1 (January 2002): 13–26.

3. Bruce Thompson, ed., *Score Reliability: Contemporary Thinking on Reliability Issues* (Newbury Park, CA: Sage, 2003); Bruce Thompson, *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications* (Washington, D.C.: American Psychological Association, 2004).

4. Colleen Cook, Fred Heath, Russell 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, 4 (August 2001): 697–706; Colleen Cook and Bruce Thompson, "Psychometric Properties of Scores from the Web-based LibQUAL+™ Study of Perceptions of Library Service Quality," *Library Trends* 49, 4 (Spring 2001): 585-604; and Bruce Thompson, Colleen Cook, and Russell L. Thompson, "Reliability and Structure of LibQUAL+™ Scores: Measuring Perceived Library Service Quality," *portal: Libraries and the Academy* 2, 1 (January 2002): 3–12.

5. Bruce Thompson and Colleen Cook, "Stability of the Reliability of LibQUAL+™ Scores: A 'Reliability Generalization' Meta-analysis Study," *Educational and Psychological Measurement* 62, 4 (August 2002): 735-43; Randall C. Arnau, Russell 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, 1 (February 2001): 23–44.

6. Catherine E. Cantrell, C.E., "Item Response Theory: Understanding the One-Parameter Rasch Model," in *Advances in Social Science Methodology*, ed. Bruce Thompson (Stamford, CT: JAI Press, 1999), 5: 171–92; David Henard, "Item Response Theory," in *Reading and Understanding More Multivariate Statistics,* eds. Larry Grimm and Paul Yarnold (Washington, D.C.: American Psychological Association, 2000), 67–98; and Robert L. McKinley and Craig Mills, "Item Response Theory: Advances in Achievement and Attitude Measurement," in *Advances in Social Science Methodology,* ed. Bruce Thompson (Greenwich, CT: JAI Press), 1: 71-135.

7. Eiji Muraki and R. Darrell Bock, *PARSCALE 4: IRT Item Analysis and Test Scoring for Rating-Scale Data* (Lincolnwood, IL: Scientific Software International, 2003).

8. Lakos and Phipps.

9. Consuella Askew-Waller, "Service Quality Evaluation Academy is a Success." *ARL Bimonthly Report*, 223 (August 2002): 11, http://www.arl.org/newsltr/223/academy.html (accessed October 6, 2004).

10. Fred Heath, Martha Kyrillidou, and Consuella Askew, eds.,"Libraries Act on Their LibQUAL+™ Findings"; Colleen Cook, guest ed.,"The Maturation of Assessment."

11. Association of Research Libraries, "LibQUAL+™ and Decision Making: The Library Summit" (Washington, D.C.: Association of Research Libraries, 2004), http://www.libqual.org/documents/admin/LibQUAL_Summit.pdf (accessed October 6, 2004).

12. Ibid.