

The Forging of Consensus: A methodological approach to service quality assessment in research libraries – the LibQUAL+™ experience

Colleen Cook and Fred Heath, Texas A&M University, USA

Martha Kyriallidou and Duane Webster, Association of Research Libraries, USA

Abstract

This paper discusses the emergence and forging of consensus in relation to the LibQUAL+ project as it developed from the winter of 1999 to summer of 2001. A rich history of assessment in the service sector using the SERVQUAL instrument over the last twenty years, a fertile ground in the library field where various institutions including Texas A&M experimented with measuring service quality at the local level for a decade, and an increasing awareness that time for action has arrived as experienced by the leaders of the ARL New Measures initiative, and members of the ARL Statistics and Measurement Committee and the ARL Leadership and Management Committee, gave birth to the historical context that created LibQUAL+. LibQUAL+ is the first total market survey developed for libraries – a standardized protocol applied across institutions that encourages libraries to complement their resource metrics with much needed metrics of users' perceived quality.

Background: The Emergence of Consensus

In the winter of 1999, North American research library leaders meeting in Tucson, Arizona, issued a call for new measures of library performance (Nitecki & Franklin, 1999). As conceived, the new measures were intended to serve as an alternative and complement to the *ARL Index*, the descriptive statistical compendium that has served the research library as virtually its only assessment metric for decades.

This paper focuses on one of the initiatives born of the Tucson symposium: LibQUAL+™, a research and development project undertaken jointly by the Association of Research Libraries (ARL) and Texas A&M University Libraries to define and measure library service quality across institutions and to create useful quality assessment tools for local planning. The emergence of LibQUAL+™ is described from the time of its endorsement by the New Measures Task Force and the ARL membership in October 1999, to the current day. The necessary building blocks of consensus are reviewed, including (1) the selection of a pilot participant group from among ARL membership, (2) the re-grounding of existent gap-theory protocol through a series of qualitative measures among the participants, (3) the development of a large-scale web-based survey device, (4) the administration of the survey and collec-

tion and analysis of the data, and (5) the dissemination of the data through publication and symposia.

Also discussed are the awarding of a \$500,000 grant from the Fund for the Improvement of Post-Secondary Education (FIPSE), a second round of implementation involving more than 40 libraries, and the scaling of the spring 2002 phase to include almost 200 participants. The paper also touches upon the plan for the transfer of the instrument from Texas A&M University to ARL for on-going administration. The challenges and opportunities that lie ahead for ARL conclude the paper.

Antecedents: Theoretical grounding of the proposed protocol

If we have learned anything from the work of Thomas Kuhn, *The Structure of Scientific Revolutions* (1962), it is that cultures and organizations are loathe to tinker with, trade, or discard explanations for complex behaviors that have served them well over time. However, when existing paradigms are increasingly unable to explain observed phenomena, emergent tools that are better able to predict and explain inevitably displace them. New ways of viewing phenomena emerge, new ways of looking at the same problems appear; the world begins to be interpreted differently.

As Danuta Nitecki has observed, on the contemporary university campus, "every unit... is valued in proportion to its contribution to the quality success of the campus" (Nitecki, 1996b). Strategically, libraries must be both competitive in today's global information marketplace, and serve as wise stewards of society's considerable investment in them. Effective service delivery must address several fundamental questions: what do users expect from libraries, and how can libraries assess whether they have fulfilled these needs? In an age of increasing accountability, academic libraries, just as other elements of higher education, are being forced to adapt themselves to the criteria by which like social institutions are commonly judged.

Fundamental to the ongoing drive to refine evaluation processes for service delivery is recognition of the sizeable investment that society makes in its libraries. For example, in the aggregate, ARL members alone expended over \$2.7 billion dollars in 1998/99 to address the library and information needs of research constituencies in North America (Kyriallidou & O'Connor, 2000). Rowena Cullen suggests yet another

motivation and exhorts libraries either to change or contemplate irrelevance: "Academic libraries are facing two major threats in the global digital environment and an increasingly competitive environment, and must improve the quality of their services in order to survive" (Cullen, 2001). She views the global digital revolution as greatly affecting not only traditional forms of information access and delivery, but also the world of higher education itself. The arrival of the virtual university with a concomitant virtual library is clearly a challenger to the *status quo*. Cullen pointedly remarks, "Retaining and growing their customer base, and focusing more energy on meeting their customers' expectations is the only way for academic libraries to survive in this volatile competitive environment" (Cullen, 2001). Today academic libraries face competition from alternative, cost-effective information providers. In light of this reality, it is imperative for libraries to seek means to ensure that their services meet and preferably exceed user expectations. A continuing program of assessment is *sine qua non* to development of a meaningful approach to meeting service goals. Again, in the words of Nitecki, "A measure of library quality based solely on collections has become obsolete" (Nitecki, 1996b).

As Sarah Pritchard explains, an assessment approach facilitating the generalizability of results across libraries and the surfacing of best practices is desirable. "The difficulty lies in trying to find a single model or set of simple indicators that can be used by different institutions, and that will compare something across large groups that is by definition only locally applicable—i.e., how well a library meets the needs of its institution. Librarians have either made do with oversimplified national data or have undertaken customized local evaluations of effectiveness, but there has not been devised an effective way to link the two" (Pritchard, 1996).

Enter SERVQUAL. "If I have seen further it is by standing on the shoulders of giants," said Isaac Newton in his famous letter to Robert Hooke. One way to respond to the pressures for change is to take a known paradigm, effectively in place in another discipline, and by careful re-grounding, reshape its tenets for application there. Librarians found one such paradigm in the commercial sector. Business leaders have long acknowledged the role of the customer in judging service. Vavra (1995) writes, "From a previous era, the ubiquitous phrase 'The customer is always right' expresses the value of customer satisfaction." From their vantage point as experts in service marketing, three collaborators – Valerie Zeithaml, A. Parasuraman and Leonard Berry – made the recognized definitive statement on the importance of the user or customer perspective: "The only criteria that count in evaluating service quality are defined by customers. Only customers judge quality; all other judgments are essentially irrelevant" (Zeithaml, Parasuraman, & Berry, 1990).

SERVQUAL, their tool for measuring service quality, has become a standard across business settings. The protocol includes 22 items measuring perceptions of service quality across the *tangibles*, *reliability*, *responsiveness*, *assurance* and *empathy* dimensions (Parasuraman, Berry & Zeithaml, 1991; Parasuraman, Zeithaml & Berry, 1985, 1994).

Danuta Nitecki and others pioneered the application of SERVQUAL to library settings (Nitecki, 1996a). Ellen Altman and Peter Hernon in their research echo the findings from the business sector, stating unequivocally that the user is the best judge of library service quality. "Many librarians maintain that only they, the professionals, have the expertise to assess the quality of library service. They assert that users cannot judge quality, users do not know what they want or need, and professional hegemony will be undermined if they kowtow to users. Such opinions about service, in fact, are irrelevant. The only thing that matters is the customers' opinions, because without users there is no need for libraries except to serve as warehouses (Altman & Hernon, 1998)."

Credibility: TAMU experience very important for the credibility of the process

As the research library community looked for new measures to respond to the pressures for accountability, a catalytic role was played by Texas A&M University, a land-grant institution in College Station, Texas. In the first place, Texas A&M was the site where Berry, Parasuraman and Zeithaml, as young assistant professors began the collaboration that gave the business world SERVQUAL. The trio continues its active collaboration, and Len Berry is still at Texas A&M University, where he holds the coveted title of distinguished professor.

Secondly, at the time the library staff at Texas A&M made its proposal to the New Measures Group at ARL in October 1999, it had already recorded six years of experience with the application of SERVQUAL to the research university environment, administering a slightly altered version of the protocol three times. Those experiences were already a part of the scholarly record (Coleman, Xiao, Bair & Chollett, 1997; Cook, Coleman & Heath 2000).

Finally, in answer to the call for New Measures, Texas A&M University Libraries was able to offer a powerful interdisciplinary research team to lead the effort to re-ground SERVQUAL and gap theory for the research library environment. In addition to library staff with previous experience with the SERVQUAL instrument, they were able to offer a strong team with a rich array of technical, quantitative and qualitative experiences. Yvonna Lincoln serves on the LibQUAL+ team as the external evaluator for qualitative methods. Dr. Lincoln is Professor and Program Director of Higher Education,

Educational Administration Department, Texas A&M University, where her leading research and teaching interests include alternative paradigm research and evaluation models, qualitative research methods, and higher education public policy questions. She is the co-author (with Egon G. Guba) of *Effective Evaluation: Improving the Usefulness of Evaluation Results Through Responsive and Naturalistic Approaches* (Jossey-Bass, 1981), *Naturalistic Inquiry* (Sage Publications, 1985), and *Fourth Generation Evaluation* (Sage, 1989); the editor of *Organizational Theory and Inquiry* (Sage, 1985); and the co-editor (with William G. Tierney) of *Representation and the Text* (SUNY Press, 1997), and (with Norman K. Denzin) of the *Handbook of Qualitative Research, 1st and 2nd Eds.* (1994, 2000). She is also the co-editor of *Qualitative Inquiry Quarterly*, a journal focused on exploring qualitative research and evaluation methodologies for illuminating public policy understandings and debate. She is also the author or co-author of over 100 journal articles, papers and chapters on evaluation and research methodology, or public policy issues in higher education. She is the former President of the Association for the Study of Higher Education, the former Vice President of Division J (Post-secondary Education) of the American Educational Research Association (AERA), and the former President of the American Evaluation Association.

Bringing equal stature to the quantitative side of the LibQUAL+ design is external evaluator Bruce Thompson. Dr. Thompson is Professor and Distinguished Research Scholar, Department of Educational Psychology, Texas A&M University, and Adjunct Professor of Community Medicine, Baylor College of Medicine (Houston). He is a Fellow of the American Psychological Association and the Editor of *Educational and Psychological Measurement* and the series, *Advances in Social Science Methodology*. He is a previous editor of two other journals. He is a past president of the Association for Assessment in Counseling and of the Southwest Educational Research Association. He is the author of 7 books, 9 book chapters, more than 160 articles, 19 editorials/notes, 10 book reviews, and 4 test reviews. Prior to his academic work, Professor Thompson was the Assistant Director of Program Evaluation Department in the Houston Independent School District. He is a licensed program evaluator (Louisiana lifetime #216). He was the National Council on Measurement in Education representative of the Joint Committee on Standards for Program Evaluation from 1992 to 1996. During this time the Joint Committee developed and published the Standards for Program Evaluation. The American National Standards Institute (ANSI) certifies these as The American National standards for program evaluation.

To the synergy of interdisciplinary campus talent could be added the contribution of CITL, the Cognition and Instructional Technologies Laboratories. Drawing upon its strengths in web-based assessment, the Texas A&M University Libraries team was able to quickly scale up the LibQUAL+ protocol for administration across the Internet. The availability of a rigorous tool placing only modest demands on local resources with no requirement for resident statistical expertise was an early factor in earning the support of the ARL community.

Collaboration: ARL's commitment to action and commitment to learning = progress

ARL's Statistics and Measurement Program is highly regarded both within the ARL community, as well as across other sectors in higher education, and other national and international programs and agencies. ARL has sustained a leadership role in the testing and application of academic library statistics in North America institutions of higher education. ARL is involved in advisory efforts for the Academic Library Statistics survey conducted on a biennial basis by the National Center of Educational Statistics (NCES); has been an active participant in efforts to define and revise standards through the National Institute of Standards Organization (NISO/ANSI); and, has been actively involved in efforts with higher education working collaboratively with the Association of American Universities (AAU), with the National Post-secondary Education Cooperative (NPEC), with the National Consortium for Continuous Improvement (NCCI), with the Association of College and Research Libraries, a division of the American Library Association (ACRL/ALA), and with the American Association of Health Sciences Library Directors (AAHSL), to name a few.

ARL sustains very close collaborative relations with its member institutions through the Visiting Program Officer program, and other arrangements that support and enhance ARL's and its member libraries' leadership role. Most notably the relation with the University of Virginia's Geosciences Data Center has made possible the electronic interactive edition of the ARL Statistics that has been widely recognized as a prototype for other agencies in developing interactive electronic publications.

Results from a recent survey of ARL member representatives to inform the efforts of the ARL Board of Directors and gauge the readiness of the current member leaders to support the ARL program framework and identify targets of opportunity show that the ARL Statistics and Measurement Program is among the top rated programs of the association. On a scale from 1-10 on how important this set of issues is to a member library, the ARL Statistics and Measurement Program

averaged 8.36 and in terms of effectiveness 7.16. In the words of some of the ARL directors responding to the survey:

“a prize program, useful on a regular, routine basis. From local benchmarking to national comparisons and a measurement-based historical record of research libraries, their successes and problems. A basis for powerful evaluations and planning – one that is getting, as it requires, a thorough review in light of the new infrastructure costs that impact libraries, the new kinds of services provided, and new kinds of tools used, the need for a determination of what really needs measuring.”

“ARL’s work allows us to describe and compare ourselves when we talk with our university leadership. In it useful to be able to say, truthfully, that ARL is working to define new measures. We need to measure ourselves in ways that show us where to invest to be more effective.”

“As they reflect the face of research librarianship, its services and its costs, these data are the most useful visible results that ARL can point to as a result of its existence.”

“Improved measures are the cornerstone for the many changes and adjustments that libraries will undoubtedly continue to require in this era of rapid change spurred by technological advances. ARL serves as a catalyst for gathering, organizing and disseminating these data.”

ARL is an active forum for collaboration where consensus emerges successfully again and again without stifling action. The building of consensus and successful collaboration is one of the ingredients of success for LibQUAL+.

Tangibility: Rapid and Visible Progress of the LibQUAL+ Initiative

One of the reputed weaknesses of collaborative efforts is the lengthy time lag that often occurs between the endorsement of an idea and the rollout of a product whose results can be shared with peers. In the not-for-profit sector, many factors may contribute to a time lag. Often the difficulty of mounting collaboration across far-flung institutions has a way of slowing the process. The costs of travel and the availability of release time of librarians from their local duties may limit agility. Further, university libraries derive their reason for being from their service to the local academic community; the earmarking of funds for developmental purposes for librarianship as a whole may carry some political costs.

In this instance of collaboration, the Texas A&M University-based initiative circumvented many of these risks, making the venture attractive to the ARL cohort. One of the reasons that the Texas A&M University

Libraries had invested in SERVQUAL locally was the dependence of library operations on a library use fee (LUF). The fee, currently at six dollars per credit hour, returns almost seven million dollars annually to the library budget. Library administrators, who took the fee to student government for review, were concerned with identifying service quality issues before they became real problems (dis-satisfiers) on their campus. Their success with SERVQUAL over a six-year period justified continued investment in a service quality assessment tool that could meet user expectations on their own campus, lead to normative information about service quality delivery, and surface best practices among research libraries.

The Texas A&M efforts through 1999 had proceeded independently of, but now fortuitously crossed paths with, the ARL New Measures Initiative. At the Tucson meeting in the winter of 1999, the participants recognized that the focus of library assessment paradigms on input measures or expenditure metrics was at variance with growing demands for evaluation and accountability. The Texas A&M proposal at the annual membership meeting in October of 1999 to underwrite a pilot project to test a new web-based tool for service quality assessment received a warm reception and quick endorsement.

The ARL-endorsed design tested score validity across libraries and across contexts, freeing institutions from sole reliance upon locally developed assessments. The emergent LibQUAL+™ project proposed to establish a national perspective, providing local managers with data fine enough to diagnose local service issues. Where deficits were discovered, managers would have the opportunity to make the improvements that best fit the local situation. At the same time, it was hoped that it might be possible through LibQUAL+™ to place the local experience in a normative context, and perhaps identify best practices across the dimensions that define library service quality. This perspective was seen as an important context for librarians and campus administrators alike (Cook, Heath, B. Thompson, & R. Thompson, 2001).

Between October 1999 and April 2000, the Texas A&M team implemented a number of steps in fast-track fashion. A diverse group of 12 ARL libraries was selected for the initial pilot phase and representatives from each of the libraries were oriented to the up-coming project at the ALA midwinter meeting in San Antonio. To ensure that the re-grounded SERVQUAL survey accurately reflected the concerns of library users at the participating universities, a series of 60 interviews took place with users (faculty, graduate students and undergraduates) over the winter months. At the same time, CITL worked with liaisons at each of the institutions to develop for each campus a customized front-end web page that would be attached to the standard questionnaire. Completing the logistical requirements prerequisite to launch, Texas A&M University

acquired the hardware and software necessary to administer the survey, capture the data, and analyze the results for a large-scale, web-based survey spanning the continent (Cook, Heath, B. Thompson, & R. Thompson, 2001).

By March 2000, less than six months after endorsement, and after beta testing with the Medical Sciences Library at Texas A&M University, the project was ready for launch. The first institutional version of the survey was loaded onto the Web on March 15, and, by the first week in June, it had run to completion on all campuses. Altogether some five thousand responses were received and automatically downloaded into SPSS for analysis. An engaged library community awaited the communications of the early findings (Cook & Heath, 2000).

Dissemination

The next step in the forging of consensus was to subject the findings of LibQUAL+™ to the scrutiny of peer review, and to answer questions about the integrity of the instrument itself as well as the desirability of its continuing application in the research library community. The first step was to share the information with the participants themselves in order to gain immediate feedback on the quality of the data and on the instrument itself. In the reporting-out meeting held for participants during ALA's annual meeting in July 2000, each of the pilot libraries was provided with mean scores for each of the questions, for each of the five first-order factors or dimensions, and an overall score. Additionally, the reporting-out session was used to assess the experiences of the pilot libraries in the administration of the survey on their home campuses. As Hendrick and Hendrick have observed, in the behavioral sciences, "theory building and construct measurement are joint bootstrap operations" (1990). The forging of consensus demands a commitment to the steady improvement of the instrument through interaction with the user community (Cook & Heath, August 2000; Cook, Heath & Thompson, October 2000).

From the outset the design team has been committed to placing the results of the research into the hands of the library community for review. Following the ALA session, the findings of the first phase were then presented at the sixty-sixth IFLA General Conference, Section on Statistics, in Jerusalem, in August 2000 (Cook, Heath & Thompson, August 2000). The LibQUAL+™ data were then explored in context of other research at an ARL-sponsored international conference on the New Culture of Assessment: Measuring Library Service Quality, in Washington, D.C., in October 2000. The papers of those well-attended sessions became the theme articles in the Spring 2001 issue of *Library Trends* (Cook & Thompson, Spring 2001; Cook & Heath, Spring 2001). With the viability of

web-based administration demonstrated, and with growing support in the library community, the design team agreed to expand the second round of participants from 24 to more than 40.

Under the guidance of their external evaluators, the Texas A&M University design team has adhered to the rigorous practice of subjecting LibQUAL+™ findings to the acid bath of peer review. In addition to a very large number of presentations at symposia and conferences, articles have been published by the research team in the *ARL Bimonthly Report*, *College and Research Libraries*, *Educational and Psychological Measurement*, the *Encyclopedia of Library and Information Science*, *IFLA Journal*, *Journal of Academic Librarianship*, *Library Information Science Research*, *Library Trends*, *Performance Measurement and Metrics*, and *portal*. An understanding of the research design, an analysis of the reliability of the instrument, and candid discussion of its strengths and weaknesses have helped to advance its acceptance in library circles.

Evaluation

Formative evaluation also takes place by analyzing direct user feedback on the LibQUAL+ survey itself. Through the analysis of user comments and due to the number of incomplete surveys we recognize that there are three issues we need to investigate further: (1) best practices when drawing random e-mail samples; (2) how to increase response rates for a web-based survey, and (3) troubleshooting and resolving software problems associated with a web-based survey. The first is to identify and establish best practices when drawing random e-mail samples. Past experience has illustrated that the quality of e-mail databases varies at each institution and we should try to provide some guidelines as to how best to draw samples from those databases.

Although results were generally representative of university populations, efforts to increase response rates will remain a high priority and efforts will be made to ascertain how response rates can be increased for a web-based survey such as LibQUAL+™. We realize that by understanding and resolving the technical problems that have accompanied our web-based survey instrument we will also likely impact the user response rate positively. Improving the technical compatibility of the instrument during fall of 2001 to diagnose browser/operating system issues and additional usability testing are the critical next step.

Since validity and reliability analyses have been conducted on a sample of well over 20,000 returned surveys from 43 diverse, mainly ARL libraries, the questionnaire has been reduced to a parsimonious base set of 25 questions, about half the length of the 2001 iteration of the survey. Thus, issues of redundancy and length, that might have affected response rate adversely, are being continuously addressed.

To gain a better understanding of what libraries expected from their participation in the pilot phase of the study, an evaluation regarding expectations and perception regarding LibQUAL+ was conducted by ARL. Three themes emerged from the responses collected. In return for their time and effort in assisting the LibQUAL+™ research team during this pilot phase of the program, participating libraries expected the following:

- To gain benchmark indicators to better assess their library service quality
- Find a means of being able to assess library service quality and provide empirical documentation of their findings to their campus administrators, and
- The development of a reliable tool by which to measure library service.

Library administrators' expectations were translated into an 18-question survey that we used to evaluate the spring 2001 LibQUAL+ administration from the perspective of the participating libraries' perceptions and expectations. The responses have been favorable (Table 1) and in all instances we have exceeded minimum expectations (see Graph 1); there is an indication that the training component of the project needs to be enhanced. Notably, several comments addressed the need for further training in deciphering their results. Specific comments to this effect include: "...Schedule more time to discuss results at participants meeting", "...establish a place on the project server to address frequently asked questions related to understanding the results (issues discussed at the participants meeting held at ALA)..." As expected, feedback surrounding technical problems, survey length and redundancy of survey questions was also provided. Overall the directors tended to evaluate the project slightly higher than coordinators (see Table 2).

Table 1: Participating Libraries' perceptions and expectation from LibQUAL+

	When it comes to...	Minimum	Desired	Perceived	Gap Between Min and Perceived
High Perceived Mean Scores regarding:					
5	adequate communication from research team members to project participants	6.69	8.56	7.80	1.11
1	understanding service quality issues and dimensions	6.33	8.49	7.44	1.11
4	providing convenient access to a reliable tool for measuring service quality	6.60	8.36	7.40	0.80
11	timely availability of the results	6.49	8.11	7.40	0.91
17	promoting a culture of assessment, i.e., a commitment to evidence-based decision making in resource allocation	6.42	8.29	7.27	0.84
3	an empirical method for identifying and measuring user expectations and perceptions	6.40	8.33	7.22	0.82
7	establishing baseline data against which you can measure the impact of future changes in service delivery	6.53	8.31	7.22	0.69
9	results that demonstrate clearly where there are service gaps in your library	6.58	8.42	7.16	0.58
2	an opportunity to contribute to the pilot LibQUAL+ process and improve it	6.16	8.09	7.11	0.96
Low Perceived Mean Scores regarding:					
16	deliverables that can be used for marketing and public relations purposes at the local level	5.84	7.71	6.60	0.76
12	establishing national norms on user expectations for library service quality	5.69	7.47	6.58	0.89
18	LibQUAL+ becoming part of your library's ongoing assessment for regional accreditation purposes	5.42	7.31	6.42	1.00
10	results that help guide resource allocation	6.13	7.98	6.38	0.24
15	external validation of previously locally conducted surveys	5.31	7.09	6.22	0.91
14	assistance/training in modifying services to better meet user expectations	5.31	7.00	5.36	0.04

Table 2: Higher average scores by the directors who evaluated the project

Descriptive Statistics Mean	Perceived			
	director	coordinator	neither	
High Perceived Mean Scores regarding:				
1	understanding service quality issues and dimensions	8.00	7.20	7.38
2	an opportunity to contribute to the pilot LibQUAL+ process and improve it	7.75	6.80	7.13
3	an empirical method for identifying and measuring user expectations and perceptions	7.92	7.00	6.88
4	providing convenient access to a reliable tool for measuring service quality	8.25	7.20	6.75
5	adequate communication from research team members to project participants	8.33	7.52	7.88
6	measuring how well current library services meet user expectations	7.58	6.40	6.38
7	establishing baseline data against which you can measure the impact of future changes in service delivery	7.92	7.00	6.88
8	obtaining informative and actionable benchmarks	7.42	6.64	7.00
9	results that demonstrate clearly where there are service gaps in your library	7.67	6.96	7.00
10	results that help guide resource allocation	6.83	6.08	6.63
11	timely availability of the results	8.50	6.88	7.38
12	establishing national norms on user expectations for library service quality	7.33	6.36	6.13
13	ability to compare your users' perceptions and expectations with those of other libraries' users	7.42	6.72	6.13
14	assistance/training in modifying services to better meet user expectations	6.58	4.84	5.13
15	external validation of previously locally	8.00	5.56	5.63
16	deliverables that can be used for marketing and public relations purposes at the local level	7.75	6.36	5.63
17	promoting a culture of assessment, i.e., a commitment to evidence-based decision making in resource allocation	8.17	6.96	6.88
18	LibQUAL+ becoming part of your library's ongoing assessment for regional accreditation purposes	7.83	6.04	5.50
		n = 12	n = 25	n = 8

In the revealing words of a director from one of the participating libraries: "The whole project more than lived up to my expectations. The empirical approach, effective teamwork and communications and the solid products make LibQUAL+ one of the most successful multi-institutional programs I have ever seen in my 35 years in higher education. Understanding the methodology, I know that LibQUAL+ cannot give us highly granular data that would answer very detailed questions like "What do users need/expect from ILL for delivery turnaround." I would ultimately like to see a method/tool that would allow us to use the more general observations that we can draw from LibQUAL+ to move to more specific questions about user needs and our performance in fulfilling them. That said - I think LibQUAL+ is a giant step in the right direction that will go down as a major watershed for ARL and other academic libraries. We expect to make it a core program in our attempts at continuous improvement" (Library Director participating in Spring 2001 LibQUAL+ administration).

External Validation

LibQUAL+™ has also won support in granting circles. In October 2000, ARL was awarded a grant from the Fund for the Improvement of Postsecondary Education (FIPSE)—a competitive and prestigious award—to continue development work on the LibQUAL+ instrument and service for three years (October 2000-September 2003). The FIPSE program is a highly selective federal program of the U.S. Department of Education. The award assures the library community that the development path of LibQUAL+ will continue, and that the mature version will be available for on-going ARL administration. The goals of the project include (a) the establishment of a library service quality assessment program at ARL; (b) the development of web-based tools for assessing library service quality; (c) the development of mechanisms and protocols for evaluating libraries; and (d) the identification of best practices in providing library service. The FIPSE funding will allow ARL to refine the questions, dimensions, and data-gathering processes and develop a cost-recovery service that ARL libraries and other academic and research libraries can use to determine their own service effectiveness. The FIPSE funds of \$498,368 will cover 49.5% of the estimated costs of the project, with ARL and Texas A&M University contributing the remaining 50.5% of the total project costs. Institutions participating in the funded project will be charged a modest administrative fee to cover direct costs of project reports (Cook, Heath & Thompson, October 2000).

Further, interest in the paradigm has extended to related fields. Texas A&M University and ARL are exploring the potential of adapting the instrument to the digital library environment and, to help support this avenue of research, are seeking funding from the

National Science Foundation (NSF) National Science, Mathematics, Engineering, and Technology Education Digital Library program. Encouragement from NSF program staff gives reason to be optimistic about the outcome of the grant application.

The attitudes of individual libraries and consortia are pivotal to the success of the LibQUAL+™. Through publications, reports, conferences, training sessions and symposia understanding of the protocol, its reliability, and its "fit" to library assessment has steadily grown. In the first phase, 12 ARL libraries were selected from among 30 volunteers as pilot libraries. In the second phase, just completed, the majority of the Big Twelve Plus Libraries (a library consortium of some 30 research universities in the western United States) and a range of other libraries participated. Altogether, 43 libraries participated, including for the first time libraries outside of the ARL community. Only 24 libraries were projected for the second iteration and 50 in the third round. The overwhelming interest and the robustness of the design have permitted a more rapid scaling up. Next year, in phase three, as many as 200 libraries may be accommodated. Among those expressing interest are the OhioLINK libraries and the American Association of Health Sciences Libraries (AAHSL). The measure of service quality across all types of post secondary libraries is an opportunity for Ohio to assess user perceptions of the investments in libraries across the state. Likewise, AAHSL contemplates the opportunity to assess service delivery across a significant proportion of that library community. The community is sufficiently committed for the National Library of Medicine (NLM) to commit to fund in part the participation fee for the AAHSL libraries.

Looking Forward: Maintenance of Consensus

Where does the library community go from here as LibQUAL+™ emerges from research and development into a full-fledged protocol administered by ARL? To this point in its development, the development team has tightly controlled the instrument, for sound psychometric reasons. Demands for the maintenance of a normative database and identification of best practices will dictate strong ARL oversight for the mature instrument. However, the pressures to accommodate the local needs of individual institutions and consortia will grow. How to balance the requirements for central control with local needs will be a policy issue that needs to be addressed in the near term.

A myriad of other questions also arise. Where normative tables are intended to provide library managers with perspective on their own scores, placing service delivery performance in the perspective of cohorts, unintended outcomes lurk in the data. With normative data, "ranking" of service quality scores becomes possible. The superior ability of one library to meet or exceed the service expectations of patrons does not necessarily

mean that library is “better” than another library when traditional metrics are concerned: investments, size of collections, etc. LibQUAL+™ is but one of many ways of listening to or assessing an organization. Continued training and communication is necessary to effect an appreciation of the instrument and an endorsement of its purposes.

The power of LibQUAL+™ and the nature of web-based surveys make possible step two activities: follow-up and longitudinal studies. Research designs incorporating those features strengthen validity testing and enable more powerful outcome measures. However, they must be balanced against increasingly rigorous requirements from institutional review boards to respect the confidentiality and anonymity of human subjects.

These policy issues underscore the need for an advisory committee. Because LibQUAL+™ has as its goal measurement of service quality across all types of North American postsecondary libraries, a diverse committee that extends beyond ARL membership could be desirable. Too, inasmuch as future designs consider the re-grounding of LibQUAL+™ for other higher education and service markets, such as digital libraries, museums, and information technology, the composition of any future advisory committee may be yet more diverse. Finally, there are issues of international applications. Philip Calvert (1997, 2000), Rowena Cullen (2000) and others have indicated the need for assessment of service delivery in all cultures, and international prospects have already been identified. International participation on the advisory committee is a reasonable expectation if consensus is to be expanded. Re-grounding of LibQUAL+™ for implementation in other cultural settings is a near-term priority.

Summary

The ingredients of consensus building are several. They include a clearly identified need, as expressed by the agenda of the New Measures Initiative and the sponsorship of a major reputable organization—in this case, the Association of Research Libraries. In the formative stage, while a critical mass builds, the sponsorship of a lead institution, underwriting essential developmental needs is essential. In this instance, the role of Texas A&M University, providing the qualitative and quantitative expertise and the large-scale web survey capability, answered a pressing need. The efficacy of rapid movement from concept, to design, to implementation and to deliverables cannot be underestimated. From the beginning, the sense of momentum was palpable and ensured and the community remained engaged in the process. Success in obtaining grant funding also played a part, providing a kind of external validation of the design project. Also important was the far-reaching dissemination of the research, evaluating the psychometric integrity of the instrument and the appropriate-

ness of its “fit” to the purposes it undertook to measure. The acceptance of manuscripts describing the instrument in some of the best library and measurement journals provided further evidence that the New Measures initiative was on track. But much work remains ahead. The instrument must be continually re-grounded for its primary clientele, and new library applications in North America and abroad must be investigated, as must its applicability to related fields. The requirements to answer the calls for greater accountability and responsiveness to user needs are greater now than they were in 1999. The challenges of consensus building, across cultures and information sectors, become even more complex.

References

- Altman, E. and Herson, P. (1998) “Service quality and customer satisfaction do matter” in *American Libraries*, 29, (7), 53-54
- Bair, L.; Chollett, B.; Coleman, V. and Xiao, Y. (1997) “Towards a TQM Paradigm: Using SERVQUAL to Measure Library Service Quality,” *College & Research Libraries*, 58, 237-251
- Berry, L.; Parasuraman, A. and Zeithaml, V. (1985) “A conceptual model of service quality and its implication for future research” in *Journal of Marketing* 49, 41-50
- Berry, L.; Parasuraman, A. and Zeithaml, V. (1994) “Alternative scales for measuring service quality: a comparative assessment based on psychometric and diagnostic criteria” in *Journal of Retailing*, 70 (3), 201-230
- Berry, L.; Parasuraman, A. and Zeithaml, V.A. (1990) *Delivering Quality Service: Balancing Customer Perceptions and Expectations*. New York: Free Press
- Berry, L.; Parasuraman, A. and Zeithaml, V. (1991) “Refinement and reassessment of the SERVQUAL scale” in *Journal of Retailing*, 67 (4), 420-450
- Calvert, P. (1998) “A different time, a different country: An instrument for measuring service quality in Singapore’s polytechnic libraries” in *Journal of Academic Librarianship* 24 (4), 296-303
- Calvert, P.J. (2000) *International variations in measuring customer expectations*. Paper presented at the Association of Research Libraries’ Symposium on the New Culture of Assessment: Measuring Service Quality. Washington, D.C. and published in *Library Trends*, 49(4) (2001)
- Coleman, V.; Cook, C. and Heath, F. (2000) “SERVQUAL: A client-based approach to developing performance indicators” in *Proceedings of the Third Northumbria International Conference on Performance Measurement in Libraries and*

Information Sciences. Newcastle-upon-Tyne, England: Information North, 211-218

Cook, C., and Heath, F. (2000) *Users' perceptions of library service quality: A "LibQUAL+™" qualitative interview study*. Paper presented at the Association of Research Libraries' Symposium on the New Culture of Assessment: Measuring Service Quality. Washington, D.C. and published in *Library Trends*, 49(4) (2001)

Cook C.; Heath, F. and Thompson. B. (2000) "A New Culture of Assessment: Preliminary Report on the ARL SERVQUAL Instrument" in *Proceedings, 66th IFLA Council and General Conference, Section on Statistics*. Jerusalem
<<http://ifla.org/IV/ifla66/papers/028-129e.htm>>

Cook, C.; Heath, F.; Thompson, B. and Thompson, R. L. (2001) "The search for new measures: The ARL "LibQUAL+™" study—a preliminary report" in *Portal: Libraries and the Academy*, 1, 103-112

Cook, C. and Thompson, B. (2000) *Psychometric properties of scores from the Web-based LibQUAL+™ study of perceptions of library service quality*. Paper presented at the Association of Research Libraries' Symposium on the New Culture of Assessment: Measuring Service Quality, Washington, D.C. and published in *Library Trends*, 49(4) (2001)

Cullen, R. (2000) Perspectives on user satisfaction surveys. Paper presented at the Association of Research Libraries' Symposium on the New Culture of Assessment: Measuring Service Quality, Washington, D.C. and published in *Library Trends*, 49(4) (2001)

Franklin, B. and Nitecki, D.A. (1999) "New Measures for Research Libraries" in *Journal of Academic Librarianship*, 25 (6), 484-487

Hendrick, C. and Hendrick, S. (1990) "A theory and method of love" in *Journal of Personality and Social Psychology*, 50, 579

Kyrillidou, M. and O'Connor, M. (2000) *ARL Statistics, 1998-99*. Washington, D.C.: Association of Research Libraries

Kuhn, Thomas (1970) *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press

Nitecki, D.A. (1996a). *An assessment of the applicability of SERVQUAL dimensions: A customer-based criteria for evaluating quality of services in an academic library*. Doctoral dissertation, University of Maryland, 1995. Dissertation Abstracts International, 56, 2918A University Microfilms No. 95-39, 711

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

Pritchard, S. M. (1996) "Determining quality in academic libraries" in *Library Trends*, 44, 572-594

Vavra, T. G. (1995) *Aftermarketing: How to keep customers for life through relationship marketing*. Chicago: Irwin Professional Publishing

Graph 1. Spring 2001 LibQUAL+ Participating Libraries Aggregate Item Summary

