

# Benchmarking Reference Desk Service in Academic Health Science Libraries: A Preliminary Survey

Kathryn Robbins and Kathleen Daniels

This preliminary study was designed to benchmark patron perceptions of reference desk services at academic health science libraries, using a standard questionnaire. Patron responses were compared to determine the library that provided the highest-quality service overall and along five service dimensions. All libraries were rated very favorably by those surveyed, but none rated significantly higher than the others except in facility appearance. Because the study revealed no other service quality differences, the results could not be used to improve services at any single library. However, the preliminary results could be useful in planning future benchmarking strategies.

Providing the best service to patrons is a worthy goal for any reference desk staff, but how can we know what constitutes the best service? Are there objective measures that can be made? Are library patrons or staff able to determine what is best service? How is “service” defined and measured, and what constitutes “the best”?

These and related questions have been asked by many librarians, and researchers have documented numerous factors that influence service quality. A 1998 publication by Peter Hernon and Ellen Altman provided thorough background information on assessing service quality in libraries using a variety of methods.<sup>1</sup> Charles A. Bunge, as well as John C. Stalker and Marjorie E. Murfin, described the Wisconsin–Ohio Reference Evaluation Program, which compares librarians’ and

patrons’ perceptions of the same reference transactions and points out the strengths and weaknesses of both sets of perceptions.<sup>2,3</sup> Jennifer Mendelsohn reported on a study designed to better understand what constitutes reference service quality by interviewing reference staff and patrons.<sup>4</sup> Other researchers have relied solely on patron perceptions and opinions, explaining that only the patrons know if they have been satisfied with a service.<sup>5,6</sup>

Although measuring service quality in a single library can be accomplished with a variety of methods, self-assessment alone may not be very useful. Comparing service quality among peer institutions can help determine whether it is the best possible, and if not, may indicate specific areas for improvement. In this preliminary study, the authors used benchmarking methodology as a means

---

*Kathryn Robbins is Head of Reference Desk Services at the Bio-Medical Library of the University of Minnesota; e-mail: krobbs@tc.umn.edu. Kathleen Daniels is a Reference Librarian and Instructor at the Minneapolis Community and Technical College Library; e-mail: danielka@mctc.mnscu.edu.*

of making reference service quality comparisons among academic health science libraries.

Benchmarking has been used in businesses for many years. It is defined as "the study of a competitor's product or business practices in order to improve the performance of one's own company."<sup>7</sup> By determining what company (library) provides the best service and finding out how that company provides services, others can adopt its best practices and thereby improve their services. Libraries have begun to use benchmarking to improve collections and services. Thomas W. Shaughnessy and Sarah M. Pritchard have provided overviews of benchmarking in libraries.<sup>8,9</sup> Recently, several reports of benchmarking studies in academic libraries have been published that look at service measurements such as wait times and Web site usage.<sup>10-12</sup> Suzanne H. Angel and Leslie G. Mackler described a benchmarking survey of hospital libraries that examined a range of library services, including reference performance.<sup>13</sup> However, the libraries that Angel and Mackler studied did not use a standardized means of determining patron satisfaction, so it is unclear how comparable their results were. Joanne G. Marshall and Holly Shipp Buchanan discussed the use of a variety of instruments to benchmark reference services.<sup>14,15</sup> One such instrument is SERVPERF, a questionnaire developed by J. Joseph Cronin and Steven A. Taylor to improve SERVQUAL, one of the classic instruments used to measure consumer perceptions of service quality.<sup>16,17</sup> Marilyn Domas White and Eileen G. Abels compared these two instruments as tools for assessing service quality and concluded that both have promise for use in special libraries.<sup>18</sup> Neither instrument has been used to survey reference services in academic libraries, although Syed Saad Andaleeb and Patience L. Simmonds used a modified version of SERVQUAL to evaluate user satisfaction with general aspects of academic library services.<sup>19</sup>

This preliminary study sought to answer the following three main questions:

1. Is it feasible to conduct a lengthy and

somewhat complex survey in a standardized manner at several geographically separated libraries?

Many elements must coincide for a multi-institutional undertaking to be successful. Administrator approval, staff support and follow-through, and patron cooperation are necessary but can be difficult to achieve, especially from a distance. The authors wanted to determine whether this could be done using only existing staff levels (i.e., without contracting with a third party to conduct the survey) and still accomplish a satisfactory outcome.

2. Do the academic health science libraries participating in this study differ in patron satisfaction with reference services?

3. If so, can these observed differences be used to improve reference services at individual libraries?

Question 2 is essentially a rewording of the null hypothesis for this study, which is that there is no difference among the participating libraries. The authors hoped to find sufficient evidence to reject the null hypothesis and then apply their findings to improve reference service, at least at the University of Minnesota. Thus, they focused their effort on measures indicating user satisfaction with existing reference services.

## Methodology

The perceived quality of reference desk service at five academic health science libraries was measured during the spring 1998 term using the SERVPERF instrument. The original wording of the questionnaire was adapted for libraries by using the term *reference desk* in place of the name of a business.

Questions 1-22 were designed to identify what aspects or dimensions of service determine the perceived quality of service. The five dimensions of service covered by the SERVPERF instrument were tangibles, reliability, responsiveness, assurance, and empathy. Sets of questions were used to measure each of these dimensions as shown in table 1. Questions

**TABLE 1**  
**Dimensions of Service Measured by the**  
**SERPERF Questionnaire**

Dimension	Definition	Questions
Tangibles	Facilities, equipment, appearance	1-4
Reliability	Perform promised services	5-9
Responsiveness	Helpful, prompt service	10-13
Assurance	Courtesy, trust, confidence	14-17
Empathy	Caring, individualized attention	18-22

Source: A. Parasuraman, Valerie A. Zeithaml, and Leonard L. Berry, "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality," *Journal of Retailing* 64 (Sept. 1988):25.

23 and 24 were intended to measure overall perceived quality and satisfaction with reference desk service. Question 25 was optional and open-ended for patrons to comment on reference services. Comments were tallied and then categorized as positive or negative.

Twelve health science libraries located at universities belonging to the Committee on Institutional Cooperation (CIC), an academic consortium including Big Ten universities and the University of Chicago, were invited to participate in the survey. Directors at six libraries agreed to participate. All six libraries served medical schools, as well as other health sciences programs, and varied in staff size (23-58 FTEs) and total annual expenditures (\$2.1 million to \$3.9 million).<sup>20</sup>

Initially, copies of the questionnaire were sent to each library along with instructions for administering the survey. Each library was asked to distribute copies of the questionnaire to patrons who requested assistance at the reference desk, until fifty completed questionnaires were returned (additional copies of the questionnaire were sent to the libraries on an as-needed basis). A systematic sampling of every other patron was to be used to permit time to explain the survey. Patron status (student, staff, faculty) was not asked. A completed questionnaire was defined as one in which (1) each question (1-24) was answered and (2) each question had only one number circled. Question 25 allowed patrons to write comments and

was optional. Each library was asked to provide a labeled box near the reference desk where patrons could return completed questionnaires. Library staff also were asked to return all questionnaires, completed or not, so that a response rate could be calculated for

each library. Distribution of the surveys ended with the close of the academic school year.

Questionnaires returned to the authors by the participating libraries were checked for completeness and the data entered into a spreadsheet. For each individual who completed all questions pertaining to a given dimension, the within-dimension responses were averaged to obtain a dimension mean. For example, if a respondent completed questions 1 through 4, the responses were averaged to obtain an in-

---

**One strength of the current study is that a standard instrument was completed during the same time period in multiple locations.**

---

dividual mean for the tangibles dimension. The individual dimension means, in addition to the responses to the overall quality and satisfaction questions, were used as outcome variables in an analysis of variance (ANOVA), conducted separately for each outcome. A variable indicating the participating library was entered as an additional independent variable. The analysis was used to determine whether any differences in dimension means or mean responses to the overall questions existed among libraries. SAS statistical software was used for all analyses (SAS Institute Inc., Version 6.12).<sup>21</sup> P-values of less than 0.05 were considered significant. No adjustment was made for multiple comparisons.

**TABLE 2**  
**Benchmarking Survey: Questionnaire Returns**

Library	Distributed	Returned	Response Rate	Complete	Partial
1	40	16	40.0	11	5
2	56	17	30.4	14	3
3	95*	48	50.5*	31	17
4	65*	51	78.5*	31	20
5	125	84	67.2	50	34

\* Two libraries did not record the numbers of questionnaires distributed. Their distribution and response rates were estimates by their library staff.

Assumptions underlying use of ANOVA include that data are randomly sampled, normally distributed, continuous variables. These assumptions were not met by the dimension means or individual questions. However, in circumstances of a reasonable sample size ( $n > 30$ ) and a lack of serious violation of the assumptions, use of these techniques is acceptable.<sup>22, 23</sup>

**Results**

Five of the participating libraries returned questionnaires. Unfortunately, only one of the libraries returned fifty completed questionnaires by the end of the survey (table 2). The return rates of two of the libraries were so low that they were eliminated from further analysis (17 and 16 questionnaires, or 30% and 40% of the number of questionnaires distributed,

respectively). The remaining three libraries returned at least 50 percent of their questionnaires.

For the three libraries remaining in the analysis, missing data for individual questions ranged from 21 percent for question 17, part of the assurance dimension, to only one percent for several questions, including 23 and 24, the overall quality and satisfaction questions. Table 3 shows mean values for the dimensions, and for questions 23 and 24, for each library.

There were no significant differences among the libraries in the overall quality or satisfaction questions (23,  $p = 0.33$ ; 24,  $p = 0.81$ ). The only dimension that differed significantly among the libraries was tangibles ( $p = 0.026$ ). A pair-wise comparison of the library means for the tangibles dimension indicated that the mean for library 5 was significantly

**TABLE 3**  
**Benchmarking Survey: Mean Values by Library**

Dimension	Library 3			Library 4			Library 5		
	<i>N</i>	<i>Mean</i>	<i>St. Dev.</i>	<i>N</i>	<i>Mean</i>	<i>St. Dev.</i>	<i>N</i>	<i>Mean</i>	<i>St. Dev.</i>
Responsiveness#	43	1.74	0.79	40	2.04	0.99	64	1.99	1.01
Empathy#	45	1.84	0.81	43	1.85	0.79	71	1.90	0.99
Tangibles*	45	6.22	0.65	47	6.16	0.76	82	5.83	1.03
Reliability*	40	6.47	0.76	41	6.40	0.73	73	6.37	1.04
Assurance*	38	6.26	0.64	39	6.30	0.70	64	6.27	0.86
Quality (23)*	48	6.35	0.67	50	6.24	0.74	83	6.43	0.75
Satisfaction (24)*	48	6.33	1.08	50	6.42	0.70	83	6.43	0.83

\* Positively worded questions; response scale 1 to 7; 7 = strongest agreement.

# Negatively worded questions; response scale 1 to 7; 1 = strongest disagreement.

lower than those for the other two libraries, whose means did not differ significantly from each other. The three libraries were not statistically different from each other on any of the other dimensions (all *p*-values were greater than 0.2). The majority of responses to the open-ended question (25) asking for comments on reference desk services were positive (73%, *n* = 70).

### Discussion and Recommendations

The failure of half of the libraries to return sufficient surveys for inclusion in the analysis was discouraging. Approval of the study by library directors, though a necessary step, was not sufficient to ensure that the surveys would be completed and returned. The authors recognize that additional groundwork must be laid in order for similar projects to be successful. Options include developing contacts among interested reference staff prior to conducting the study, with continued communication throughout its course. Without at least one person at each library committed to completing the survey, it is unlikely that response rates will improve. A more costly alternative would be to have the survey administered by a third party. This would present logistical problems if nonlibrary personnel were needed at the reference desk to distribute questionnaires.

One strength of the current study is that a standard instrument was completed during the same time period in multiple locations. This should improve our ability to make comparisons and make us more confident in their results. However, as an alternative, if a standard instrument was used by different libraries on an ongoing basis and the results made available to other institutions (as has been suggested for Web site statistics<sup>24</sup>), an individual library could use the previously collected results for benchmarking.

The majority of user responses from all libraries included in the analysis indicated high satisfaction with reference desk services. The average response for the three libraries combined was 6.4 on

a scale of 1 to 7 for the satisfaction question. The comments written in response to optional question 25 further substantiated this overall positive assessment.

Unfortunately, this nearly unanimous high approval cannot be used to benchmark service, as originally intended, because the libraries did not differ. The survey failed to identify a single "best" library on either of the overall measures as well as on four of the five service dimensions, making the results inadequate for benchmarking purposes. The one area in which the libraries differed was tangibles, which includes physical facilities, equipment, and appearance. Interestingly, the lowest-rated library has since been re-modeled.

The methodology used in the study was sensitive enough to pick up the difference among the libraries in physical appearance. Why, then, was no difference seen on any other measure? Explanations include the possibility that all the libraries provide uniformly high-quality reference service, with highly satisfied patrons. This possibility cannot be ruled out, on the basis of the current results (nor do the authors particularly want to rule it out), but other possibilities come to mind. It is possible that only satisfied users, those with a strong motivation to "help out" the library, agreed to complete the questionnaire or took the time to complete and return it. In addition, it is possible that, subconsciously, those handing out the questionnaire were more likely to give it to patrons who appeared to be agreeable, despite instructions to give one to every other patron. Finally, it is possible that unhappy patrons are less likely to even enter the library. All of these sources of potential bias could be ameliorated by taking a randomized sample of potential reference service users, including those who physically enter the library, those who use only electronic resources, and those who may not use the library at all. Of course, this also complicates the logistics of conducting a survey but could serve to provide a broader range of opinions regarding reference services.

Notes

1. Peter Hernon and Ellen Altman, *Assessing Service Quality: Satisfying the Expectations of Library Customers* (Chicago: ALA, 1998).
2. Charles A. Bunge, "Gathering and Using Patron and Librarian Perceptions of Question-Answering Success," in *Evaluation of Public Services and Public Services Personnel*, ed. B. Allen (Champaign, Ill.: University of Illinois Graduate School of Library and Information Science, 1991), 59–83.
3. John C. Stalker and Marjorie E. Murfin, "Quality Reference Service: A Preliminary Case Study," *Journal of Academic Librarianship* 22 (Nov. 1996): 423–29.
4. Jennifer Mendelsohn, "Perspectives on Quality of Reference Service in an Academic Library: A Qualitative Study," *RQ* 36 (summer 1997): 544–57.
5. Wendall Sullivan, Lisa A. Schoppmann, and Patricia M. Redman, "Analysis of the Use of Reference Services in an Academic Health Sciences Library," *Medical Reference Services Quarterly* 13 (spring 1994): 35–55.
6. David W. Harless and Frank R. Allen, "Using the Contingent Valuation Method to Measure Patron Benefits of Reference Desk Service in an Academic Library," *College & Research Libraries* 60 (Jan. 1999): 56–69.
7. WWWebster Dictionary. Available online at: <http://www.m-w.com>.
8. Thomas W. Shaughnessy, "Benchmarking, Total Quality Management, and Libraries," *Library Administration & Management* 7 (winter 1993): 7–12.
9. Sarah M. Pritchard, "Library Benchmarking: Old Wine in New Bottles?" *Journal of Academic Librarianship* 21 (Nov. 1995): 491–95.
10. Margaret Robertson and Isabella Trahn, "Benchmarking Academic Libraries: An Australian Case Study," *Australian Academic and Research Libraries* 28 (June 1997): 126–41.
11. Joy Tillotson, Janice Adlington, and Cynthia Holt, "Benchmarking Waiting Times," *College & Research Library News* 10 (Nov. 1997): 693–94, 700.
12. Christy Hightower, Julie Sih, and Adam Tilghman, "Recommendations for Benchmarking Web Site Usage among Academic Libraries," *College & Research Libraries* 59 (Jan. 1998): 61–79.
13. Suzanne H. Angel and Leslie G. Mackler, "A Benchmark Instrument Tested in Women's Hospital Libraries," *Bulletin of the Medical Library Association* 84 (Oct. 1996): 582–85.
14. Joanne G. Marshall and Holly Shipp Buchanan, "Benchmarking Reference Services: An Introduction," *Medical Reference Services Quarterly* 14 (fall 1995): 59–73.
15. Holly Shipp Buchanan and Joanne G. Marshall, "Benchmarking Reference Services: Step by Step," *Medical Reference Services Quarterly* 15 (Sept. 1996): 1–13.
16. J. Joseph Cronin and Steven A. Taylor, "Measuring Service Quality: A Reexamination and Extension," *Journal of Marketing* 56 (July 1992): 55–68.
17. A. Parasuraman, Valerie A. Zeithaml, and Leonard L. Berry, "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality," *Journal of Retailing* 64 (Sept. 1988): 12–40.
18. Marilyn Domas White and Eileen G. Abels, "Measuring Service Quality in Special Libraries: Lessons from Service Marketing," *Special Libraries* 86 (winter 1995): 36–45.
19. Syed Saad Andaleeb and Patience L. Simmonds, "Explaining User Satisfaction with Academic Libraries: Strategic Implications," *College & Research Libraries* 59 (Mar. 1998): 156–67.
20. Association of Academic Health Sciences Libraries, *Annual Statistics of Medical School Libraries in the United States and Canada*, 21st ed. (Seattle, Wash.: Association of Academic Health Sciences Libraries, 1999).
21. SAS Institute, Inc., *SAS/STAT User's Guide, Version 6*, 4th ed., Vol. 2 (Cary, N.C.: The Institute, 1990), 891–996.
22. David B. Allison, Bernard S. Gorman, and Louis H. Primavera, "Some of the Most Common Questions Asked of Statistical Consultants: Our Favorite Responses and Recommended Readings," *Genetic, Social, & General Psychology Monographs* 119 (May 1993): 153–85.
23. George W. Snedecor and William G. Cochran, *Statistical Methods*, 7th ed. (Ames, Iowa: Iowa State University Pr., 1980), 204–6.
24. Hightower, Sih, and Tilghman, "Recommendations for Benchmarking Web Site Usage Among Academic Libraries," 77.