

Reaching Reference Service Excellence: Developing a Mystery Shopping Program to Measure Service Quality, Performance, and the Patron Experience

Elizabeth Kocevar-Weidinger and Candice Benjes-Small

Introduction

In the current atmosphere of high-stakes accountability, how can library managers most accurately measure the intangible product of public service staff transactions with patrons? Since the inception of reference/user services by Swet and Melville Dewey in the late 19th century, there has been interest in how accurately and effectively library staff delivers service.¹ Past studies have made accuracy, customer service, and user satisfaction the benchmarks for library service points. As the authors of this article contemplated finding the most effective ways to measure these benchmarks, we both recalled our former lives in the retail world and how much importance was placed on quality service. Both of us had worked in stores that used mystery shopping to measure service successfully. In a mystery shopping exercise, a proxy acts as a shopper and then rates his/her experience, based on criteria provided by the employer. As librarians, we reflected on our past experiences, expanded our research beyond the library literature, and delved into the business literature to determine if mystery shopping would work in our current environments. Verifying that mystery shopping is often used to measure both customer service

and adherence to policies, we concluded that mystery shopping might be a viable method to measure the quality of service at library reference and circulation desks.

Reference Service Excellence Literature Review

Although measuring reference service has been a long-time goal of the profession, renewed interest began to truly mount following Crowley and Childers' book, *Information Services in Public Libraries: the Studies*, published in 1971.² This work launched an emphasis on unobtrusive evaluation techniques, in which the library staff does not know they are being studied or evaluated. The most famous of these studies completed in the library world is Heron and McClure's 1986 article, "Unobtrusive Reference Testing: The 55 Percent Rule." This rule, often whispered about in library hallways, concludes that library staff answered questions accurately only 55% of the time.³ In her bibliographic essay reviewing recent research on reference effectiveness, Emily Rimland (2007) reports that numerous studies followed, trying to disprove Heron and McClure's conclusion.⁴ More recently, li-

Elizabeth Kocevar-Weidinger, e-mail: kocevarweidingerea@longwood.edu; Candice Benjes-Small, e-mail: cbsmall@radford.edu.

brarians focused on a more holistic view of the user's experience; being treated well by library staff was just as important as getting the right answers. Librarians conducted studies to examine whether the user would voluntarily return to the desk for follow-up questions,⁵ how reference transactions ended,⁶ and how the nonverbal communication of librarians affected users.⁷ In Pali Kuruppu's article, "Evaluation of Reference Services—A Review" she posits that "although the amount of evaluation research has increased with time, the best measures to use in evaluation of reference services continue to be uncertain and elusive;"⁸ confirming our own experiences in this area.

We shifted our focus toward exploring the user's total experience at the library reference desk; a large component was the manner in which the customer received service. Seeking to measure service at our libraries' reference desks, we read through the existing literature in our field. We were intrigued by the aforementioned studies, but also wondered whether we might be able to borrow techniques from the business world in which customer service is often life or death to an organization.

Service Excellence in Business

Few would argue that customer service is imperative for a business to survive.⁹ As in the library world, studies in business publications showed that customer service is not just about the right answer; equally important is the sense that the customer service representative cares. The business literature often refers back to a 1991 study by Tucker and Gulbro which found that 2/3 of bank customers changed banks due to dissatisfaction with an employee's attitude of indifference, confirming the saying, "customers don't care what you know until they know that you care."¹⁰

While businesses may use numerous tools to measure customer service, one that is noted as being particularly effective is mystery shopping. Mystery shopping can be used to judge an employee's performance, confirm adherence to policies, and measure the effectiveness of training programs. Research supporting the effectiveness of mystery shopping has been published extensively, particularly in retail, tourism, and banking.¹¹ Interestingly, one of the articles we found most informative, Kate Purcell's "Mystery Shopping in the Banking Industry,"¹² was written by a librarian and published in a library publication.

Based on this literature, the authors concluded that mystery shopping is a viable way to measure service excellence and could be applied to a library environment.

Mystery Shopping in Libraries

The authors decided to implement cyclical mystery shopping exercises at their libraries to gauge customer service at the reference desks. The authors hoped that by repeating the exercises regularly, they could chart growth and identify areas for improvement over time. Although the library literature includes articles about mystery shopping as used in libraries,¹³⁻¹⁵ none focused on using the exercises in an iterative way or sought out direct expertise of the business sector. Instead, we chose to use the business literature to help us construct the exercises, and then use a marketing professor's expert advice to create a framework which could be used each year. Designing a sustainable program created some challenges, but we learned more with each exercise.

Developing Guidelines and Buy-in

Mystery shopping should be marketed as part of a process, not an end in itself. Our goal was to provide consistent, exemplary service at the library desk, not score well on a "test" of some sort. Viewing mystery shopping as a test may frighten or threaten staff members and make them unwilling/reluctant to participate. Therefore, rather than springing a mystery shopping exercise upon staff, we believed it was critical to involve them as co-creators of the guidelines and therefore, the assessment measure. We assured staff that we would all be assessed, and that we would all be involved in conversations about the results. The mystery shop would be part of a cycle of service creation, implementation, assessment, change and innovation (rather than improvement).

At our libraries, our staff revealed that while good customer service was an objective embraced by all, there was little consensus as to what amounted to the "ideal" service. As outlined by Kate Purcell, we agreed that "the aim of a mystery shopping exercise is performance measurement and service evaluation based against a set of predefined standards."¹⁶ Each institution conducted customer service workshops in which the staff worked together to develop guidelines for what they believed was "good" customer service. The staff were provided with resources such as the Refer-

ence and User Services Association's Guidelines for Behavioral Performance of Reference and Information Service Providers¹⁷ and worked as a team to develop their department's version. These sessions were designed to generate customer service guidelines, generate staff buy-in and make the process transparent. After the guidelines were developed, we held discussions to determine if the guidelines were reasonably obtainable on a regular basis. We addressed roadblocks to delivering quality service, and came to consensus that the guidelines were reasonable in the current environment.

Like Purcell, we believed that "the exercise must have the support and confidence of front-line staff... and to be ethical; the staff must be advised that their service delivery may be checked."¹⁸ Therefore, from the beginning, the staff knew that they would be evaluated by their self-created guidelines. The guidelines were distributed to all students and staff who worked at the reference desks, and were placed online. Staff participated in training sessions on the guidelines as well as on reference interviewing. Once we established that mystery shopping would be the assessment method, we made the process as transparent as possible. We led sessions to inform staff about what mystery shopping was (and was not); how we would implement the process; who the data would be shared with; how the data would be used, and provided opportunities for input and comments. We each obtained campus IRB approval for our assessment, and had all participants complete informed consent forms when required.

To alleviate any initial apprehension, we told staff the dates of the first mystery shopping exercise. We believed this would quell the anxiety of the evaluation process. In hindsight, we realized that this somewhat invalidated our results—staff knew when to be on "best behavior." After a very positive first test-run, and staff agreed there was a moderate level of comfort, dates were not provided for future exercises.

We constructed the evaluation forms that mystery shoppers would use to record their "shopping" experiences using the customer service guidelines developed by the staff and a pool of frequently-asked reference questions developed by desk supervisors. We chose the reference questions that we believed would elicit the service behaviors being tested. The actual questions were modified so that even the supervisors did not know which questions were being asked by mystery shoppers. To keep the shoppers' workload man-

ageable, only the major behaviors were tested. At the suggestion of our marketing expert, we developed a 3-point range for each behavior. For example, for the question, "Were you greeted promptly?", possible answers were, "I was not greeted;" "I was greeted after I said hi or hello;" "Employee immediately greeted me." In choosing answers, we wanted actions which reflected undesirable, adequate, and desirable behaviors. The customer service evaluation form also included space for the shopper to indicate what answer s/he was given in response to the question and a qualitative, open-ended comment option. The students' comments provided some of our most valuable insights into the students' service expectations and our service delivery.

Implementing the Exercise—The Business Way

Collaborating with business faculty in each respective institution, we conducted an initial in-house mystery shopping assessment, much like Hernon and McClure's work, but enhanced with business model expertise. The librarians also consulted, Dr. Angela Stanton, a marketing professor at Radford University with experience running mystery shopping studies for businesses, to discuss and review our model to ensure that it met marketing standards.

Librarians and marketing faculty at each institution worked together to train students for their role as mystery shoppers. Each of the faculty members volunteered to solicit one or more of their classes to participate in the exercise. It was the faculty member's decision whether or not to offer extra-credit or a grade as an incentive for students to volunteer to shop the library staff. We knew that our business faculty would be the knowledgeable collaborators and perhaps the most cooperative in this project. We hoped that the business students would have the business angle interest and awareness of the make-or-break nature of customer service delivery.

The librarians came to the faculty member's classes to lead the training. Students were informed about what mystery shopping was, why it was important, and what the service guidelines for service were for each library. Through presentation and role playing, we trained students to be familiar with the library's service guidelines, to memorize and ask scripted questions as a patron, and how to complete evaluation forms immediately after shopping the staff. Our goal

was to have shoppers who were knowledgeable about the kind of service they should receive, who would appear to be “average” patrons, and who could provide the information we sought via the evaluation form.

Assessment and Analysis

Our mystery shopping exercises ran for two weeks during the semester. Each desk was visited approximately 20 times. (After the initial exercises, staff did not know which two weeks, just that the exercise was being conducted sometime in the semester.) The faculty members collected the shoppers’ reports and gave them to us. A designated librarian at each library generated a final report which varied slightly for each institution. These reports, which included answers to questions, comments about the service, and other qualitative data were shared with the desk supervisors, the library administration, and most importantly, the staff who had been shopped. We expect to analyze this data as part of an iterative assessment process that allows us to track and modify behaviors over the course of time.

Through the mystery shopper program, we have gained a picture of what level of service is generally found at our service desks and if specific types of questions were answered correctly. Surveys like LibQual+ often provide comments from users about customer service, but such comments are usually from people who have had a very memorable experience (either good or bad), rely on patrons’ memories of past experience rather than immediate experience, and are not customized to local guidelines and therefore may not provide a complete picture. Through mystery shopping, we had evaluations of multiple visits within a short time period; the shoppers saw different employees, at different times of day. Although some administrators and staff predicted we would discover deficiencies in our service, in all cases we found much to celebrate. Overall, the shoppers were very pleased with the service provided, including the manner in which we interacted with the patron AND the correctness of the answer. Of course, the exercises’ results also identified areas for improvement. Employees may have developed and been trained to follow guidelines, but that did not guarantee adherence. Desk supervisors now had data to develop appropriate training or revamp guidelines for realistic goals, then use mystery shopping again to see if these new strategies are viable. Thus, the circle of assessment is complete.

Conclusions

Quality customer service is essential at library service desks. Like Van der Wiele’s article in which he concludes that the “mystery shopping instrument is useful for improving service delivery,”¹⁹ we believe that our methodology yielded productive results upon which we could build better customer service. Also, like Van der Wiele and Calvert, we believe that mystery shopping is one of many methods to measure what we do and how well we do it. Measuring customer perceptions of service is a critical component in the cycle of learning what can be done to improve service quality. Literature in the business and library worlds indicated that mystery shopping is a valid, business-tested method upon which librarians can develop assessment programs. Managers who wish to thoughtfully develop and skillfully implement mystery shopping programs may want to consider the following issues:

1. staff buy-in,
2. solid guidelines which translate into benchmarks,
3. transparency,
4. shopper training,
5. consistent implementation, and
6. meaningful action on the results for improved services.
7. mystery shopping to be conducted regularly, so as to chart progress and identify weaknesses in customer service.

Measuring customer perceptions of service is a critical component in the cyclical evaluation process. We believe that mystery shopping is an effective, meaningful assessment tool in the iterative process of measuring service.

Notes

1. Wayne A. Wiegand and Donald G. Davis, *Encyclopedia of Library History* (New York: Garland Publishers, 1994), 541-542.
2. Terence Crowley and Thomas Childers, *Information Services in Public Libraries: The Studies* (Metuchen, NJ: Scarecrow Press, 1971).
3. Peter Hernon and Charles R. McClure, “Unobtrusive Reference Testing: The 55 Percent Rule,” *Library Journal* 111 (1986): 37-42.
4. Emily Rimland, “Do We Do It Good Well? A Bibliographic Essay on the Evaluation of Reference Effectiveness,” *The Reference Librarian* 47: 44.
5. Joan C. Durrance, “Reference Success: Does the 55

Percent Rule Tell the Whole Truth?" *Library Journal* 114 (1989): 31-36.

6. Patricia Dewdney and Catherine S. Ross, "Flying a Light Aircraft: Reference Service Evaluation from a User Viewpoint," *RQ* 34 (1994): 217-231.

7. Marie L. Radford, "Communication Theory Applied to Reference Encounter," *Library Quarterly* 66(1998): 123-138.

8. Pali U. Kuruppu, "Evaluation of Reference Services-A Review," *The Journal of Academic Librarianship* 33(2007): 369.

9. C. Erlich, "Mystery Shopping," *Competitive Intelligence Magazine* 10(2007): 43-44.

10. Randy Spitzer, "The Case Against Targets!" *The Journal for Quality and Participation* 28 (2005): 30.

11. Ton Van Der Wiele, Martin Hesselink, and Jos Van Iwaarden, "Mystery Shopping: A Tool to Develop Insight into Customer Service Provision," *Total Quality Management* 16(2005): 529-541.

12. Kate Purcell, "Mystery Shopping in the Banking Industry," *Library & Information Update* 4 (2005): 38-39.

13. Mary Tygett, V. Lonnie Lawson, and Kathleen Weessies, "Using Undergraduate Marketing Students in an Unobtrusive Reference Evaluation," *RQ* 36 (1996): 270-275.

14. Philip Calvert, "It's a Mystery: Mystery Shopping in New Zealand's Public Libraries," *Library Review* 51(2005): 24-35.

15. Merrill Stein, et. al, "Using Continuous Quality Improvement Methods to Evaluate Library Service Points," *Reference & User Services Quarterly* 48(2008): 78-85.

16. Purcell, "Mystery," 38.

17. Reference and User Services Association, "Guidelines for Behavioral Performance of Reference and Information Service Providers," <http://www.ala.org/ala/mgrps/divs/rusa/resources/guidelines/guidelinesbehavioral.cfm>.

18. Purcell, "Mystery," 38.

19. Van der Wiele, "Mystery Shopping," 540.