# Quick and Easy Reference Evaluation
## *Gathering Users' and Providers' Perspectives*

**Jonathan Miller, Guest Columnist**

*Corresponding concerning this column should be addressed to* ***Judith M. Nixon****, Head, Humanities, Social Science & Education Library, 504 West State Street STEW, Purdue University, West Lafayette, IN 47907; e-mail: jnixon@purdue.edu.*

***Jonathan Miller*** *is Director of Libraries, Olin Library, Rollins College, Winter Park, Florida.*

Imagine a reference survey instrument that is very simple to administer, requires only a pencil to fill out, and gathers data specifically on *whether users get the help they need and are satisfied with reference service, and whether in the process they learn about how to find and evaluate information*. If you are interested, read on. This third article in the new Management column is written by Jonathan Miller about the reference survey he and his colleagues developed at the University of Pittsburgh. I first met and heard Miller at the 2006 American Library Association (ALA) Annual Conference when he presented his research at the Reference Research Forum. I found it a fascinating and practical approach to obtaining user input; a survey that combined some of the strengths of the Wisconsin Ohio Reference Evaluation Project (WOREP) survey while overcoming some of the shortcomings of that standard workhorse instrument. I was especially interested because this new survey was developed to build upon the survey data libraries gathered from LibQUAL+. LibQUAL+ does not specifically measure reference quality; this survey provides a way of gathering useful evaluation of reference service.

A future column will be about a mid-life librarian who changed careers from a special librarian to an academic librarian, including a discussion of the ups and downs of such a change and the energizing effect of the change. This is of timely interest to administrators and managers because within the next few years a large number of librarians will retire, opening up many higher level positions. Mid-career librarians will be in the prime position to apply for these jobs, if they are willing to make some career changes.

For future columns I am looking for articles on new approaches to managing reference service, such as outreach services in non-library locations, or the use of expert systems, or using instant messaging (IM) as a way of communicating with high school or college students. If you have experiences with any of these or other practical ideas for providing reference service, please e-mail me.—*Editor*

**This research grew from my concern as a public services librarian-manager to find a quick and** easy way to evaluate reference service. Most, perhaps all, libraries measure how much reference assistance we provide. Usually we simply count the number of transactions, sometimes classifying these transactions by complexity (directional, ready reference, and so on) or in terms of medium (in-person, phone, and online). Sometimes we even collect information on how long it takes to complete the transaction. As a profession we have developed a variety of definitions of what constitutes a reference transaction to help with the collection of data.[1] RUSA has developed reference behaviors guidelines that I have found to be very useful during training and development of reference providers (my preferred, if awkward, term for all those employees—librarians, staff, interns, and student employees—who may provide reference service).[2] It is more difficult to measure the quality of the reference service we provide, particularly if we want to avoid overburdening our reference providers and our users. The LibQUAL survey of library service quality asks about "employees who have the knowledge to answer user questions," "employees who deal with users in a caring fashion," and "employees who understand the needs of their users."[3] All of these statements could be related to reference providers, but LibQUAL does not single out individual library services, such as reference. Instead, quite rightly, it asks for users' perceptions of library service quality as provided by the whole library. These perceptions could be formed on the basis of interactions with any library personnel with whom the user interacts.

## REFERENCE EVALUATION LITERATURE

The literature on reference evaluation is huge, and this column is not the place to review it. But two works in particular helped me understand some of the underlying issues of reference evaluation and amply repay review by anyone interested in evaluating reference service. Saxton and Richardson's 2002 book is an excellent evaluation study.[4] They also critically evaluate earlier reference research. Jo Bell Whitlatch's 2000 book provides a very thorough practical guide to reference evaluation.[5]

The research that is most relevant to the work presented here combines the users' and the providers' assessment of the reference transaction. The idea of gathering data from both the user and the provider has been used by a number of researchers, including Whitlatch in 1990.[6] The most prominent research in this area is the Wisconsin Ohio Reference Evaluation Project (WOREP.) The Reference Transaction Assessment Instrument (RTAI) that grew out of WOREP was developed by Charles Bunge, Marjorie Murfin, and Gary M. Gugelchuk.[7] WOREP and the RTAI are currently available at http://worep.library.kent.edu. The great benefit of the RTAI is that it is standardized, it is statistically valid and reliable, the results are machine readable (thus saving time and resources on data entry), and the results can be compared to other reference service units. But the instrument itself is daunting, especially from the provider's perspective. As one colleague put it upon seeing a copy, "it looks like the SAT." It takes some minutes to complete (minutes that the provider might not have and that the user may be unwilling to invest), and asks for a level of detail about the interaction that may be greater than the user in particular is able to provide. The challenge that we faced was to combine the power of evaluations from both sides of the reference desk in evaluating reference service and yet not overwhelm providers or users of the service.

## DEVELOPING A METHOD

Rather than developing an evaluation method at a single reference desk, we brought together a variety of academic library reference operations in Pittsburgh, including public and private, large and small, and libraries that support a wide range of users working in many subject areas and at many levels. Thus the questionnaire we developed has been tested in a variety of settings, all academic, but representing different libraries with varying models of reference service and users. The specific libraries involved were the Hillman Library Information Desk and the Engineering Library at the University of Pittsburgh, the Hunt Library, Engineering and Science Library, and Arts Library of Carnegie Mellon University, and the Chatham College Library. We formed a working group of representatives from each library to develop and test the instrument and analyze the results.

The first step was to agree on the outcomes of reference service. Since our aim was to develop an easy and quick way of measuring the quality of our service, our outcomes were necessarily simple and general. They were certainly not detailed or specific enough to meet the requirements of formal social scientific research. They were:

1. The user gets the information they need.
2. The user learns something about how to find information.
3. The user learns something about how to evaluate information.
4. The user is satisfied with the interaction.

We decided to study the perceptions of users and providers of reference services with all the limitations that perception studies imply. The model we used to explain to people what we wanted to do was the simple customer satisfaction postcard used at restaurants, but in our case a postcard would be completed by both the customer and the server.

We developed a short questionnaire (see figure 1), tested it, and made some minor adjustments to wording. The questionnaire is printed here so any library can modify and use it. The two-part form is perforated (using a simple manual paper perforator available from any office supply company) and is separated when handed to the user. Each copy of the questionnaire includes a control number on each part so responses can be linked during data analysis (that is, the same number appears on the top [user] and bottom [provider] portion of the form. A different number appears on each sheet.) This was achieved with a simple mail merge in Microsoft Word. The use of a sequential control number has the added advantage of enabling us to calculate how many surveys had been distributed at any one service point and thus calculate the response rate. The questionnaire asks some basic demographic information about both user and provider and also provides an opportunity for users to make comments and to identify themselves if they wish to be included in future, perhaps qualitative, assessment of reference service (for example, focus groups and interviews) The core of the instrument is the paired questions that seek the user's and the provider's perspective on whether each of the outcomes were met.

## TESTING THE INSTRUMENT

In our test of the questionnaire, we could have spread the sample out, handing the form to every tenth user, for instance, but we decided that this would be overly complicated and confusing for reference providers. So we simply administered the survey for a three week period during our busiest times (in the middle of the fall semester). Some participant libraries in the test were not interested in the statistical significance or in generalizing the results of our test. But it is possible to calculate samples based either on the total population served by any library or on the total number of reference questions asked at a service point. For instance, working with a statistician at the University of Pittsburgh, we calculated an appropriate sample size that was based on both the total number of reference transactions at the Hillman Library and on the number of faculty and students at the university. In effect we considered two possible definitions of the population: the total number of reference transactions, or the total number of potential users of reference. In both cases the sample sizes required to have some confidence in the representativeness of our results were small. To achieve results within a 5 percent margin of error at (95 percent confidence) we would need responses from only thirteen University of Pittsburgh undergraduates, graduates, and faculty. In this case it was necessary to assume that who asked questions (undergraduate, graduate, and faculty) and also the subject and complexity of those questions was random over time.

Most of our outcomes, except perhaps for outcomes 1 and 4, are not relevant to directional questions, such as "where are the bathrooms?" We did not want to ask users and providers to complete a questionnaire each time they asked such directional questions for which reference services are not designed and organized. Although reference providers certainly answer many such questions because these often lead to more substantial questions, and answering them reinforces the important lesson that reference providers are here to help; they are not the *raison d'etre* of reference service. To exclude such questions, which form forty-three percent of the total transactions recorded at the main Hillman information desk, and to ensure consistency between libraries, we created the following rubric to help the reference providers decide when to offer the questionnaire to users. We asked them to only offer the questionnaire if the interaction meets the following criteria:

> The question or need for service is greater than an answer to a simple directional question or a request for a simple query of the library catalog or database to locate a known item. As a general rule, if it takes longer to complete the questionnaire than it did to complete the transaction, don't ask them to do it.

We developed this statement after discussion with the reference providers who would have to administer the survey. We also gave providers some form of words they could use to introduce the questionnaire to the user and to meet the requirements of our institutional review boards.

> We are evaluating our service this week. Can I ask you to fill out this short survey and put it in that box? [Point to box.] It is voluntary and your answers are anonymous. If you have completed one before, your answers are still useful since the questions concern this single transaction. Thanks.

**Figure 1:** Questionnaire used by Pittsburgh libraries

**Status (check one)**
❑ [school] Undergrad. ❑ [school] Grad. ❑ [school] Faculty
❑ [school] Staff ❑ Other

Please rate the following statements by circling a number (1 is strongly disagree, 5 is strongly agree) or not applicable (N/A).

I got the information I needed.
1 2 3 4 5 N/A

I have a better understanding of how to find information.
1 2 3 4 5 N/A

I have a better understanding of how to evaluate information.
1 2 3 4 5 N/A

I am satisfied with the service I received.
1 2 3 4 5

Yes! I am interested in participating in further evaluation of library services. Here is my e-mail address: _________@_____________

*If you have comments, please write them on the back of this card. Please place the completed survey in the box by the exit. Thank you.*

---

**Status (check one)**
❑ Librarian ❑ Staff ❑ Student

Please rate the following statements by circling a number (1 is strongly disagree, 5 is strongly agree) or not applicable (N/A).

The user got the information he or she needed.
1 2 3 4 5 N/A

The user has a better understanding of how to find information.
1 2 3 4 5 N/A

The user has a better understanding of how to evaluate information.
1 2 3 4 5 N/A

Overall, I was satisfied with this reference transaction.
1 2 3 4 5

*If you have general comments, please write them on the back of this card. Thank you.*

We printed more than the number of questionnaires we estimated might be needed at each service point and distributed them to all service points. We reminded all reference providers that the study would be conducted for the specified period and asked them to hand the top part of the form to the user at the end of the reference transaction. We also asked them to complete the bottom portion as soon as possible after the transaction and deposit the completed questionnaire in a box located conveniently at or near the service point. At the end of the survey period, the responses were collected from both users and providers and entered into a Microsoft Excel spreadsheet. Each row consisted of the data from surveys with a single control number, bringing together the responses of both the user and the provider to a single reference interaction for analysis.

## RESULTS

The purpose of this column is not to report on a particular research study but to introduce a method for evaluating reference service that captures both the user's and the provider's perspective. If the reader is interested in a full report on our use of the instrument in Pittsburgh academic libraries, the PowerPoint presentation to the Reference Research Forum at the 2006 ALA Annual Conference can be found at http://web.rollins.edu/~jxmiller/Refevaluation.htm.

This questionnaire measures user and provider perceptions of the success of individual reference transactions as measured by whether users received the information they needed, learned something about how to find information, learned how to evaluate information, and whether both parties were satisfied with the reference interaction. With enough responses it is possible to drill down into the results to explore distinctions between the perceptions of various groups of users or providers (in our test, for instance, the mean responses for those transactions involving a librarian were consistently higher than those for staff) and also, with appropriate statistical software, to be able to explore the relationships between the data. For instance, do undergraduate students perceive that they are learning how to evaluate information during the reference interaction? Or does the provider's perception of a successful reference interaction correspond to the user's perception of the same interaction?

One example of what can be learned from the questionnaire concerns the role of information evaluation in reference. The lowest mean results we received in our test were for questions three and seven, concerning the outcome "the user learns something about how to evaluate information," and that was also the question with the lowest or number of paired responses. This result reflects the fact that we got the highest number of "not applicable" responses to this question (and these came largely from providers, not users.) We can interpret this result to mean that our users perceived that they were learning how to evaluate information during reference transactions, even when our reference providers did not think that any teaching, learning, or modeling of information evaluation was taking place. Or if such instruction was taking place, providers did not perceive that users were learning how to evaluate information. Evaluation is a major component of information literacy, and as such is a major part of what we do at the academic reference desk. In this case our users seemed to recognize that, but our providers did not. Perhaps the lower mean was a result of that lack of conscious attention to the issue of evaluation by providers. This result was an opportunity to discuss the role of information evaluation during the reference interaction and to find ways to improve our ability to explicitly raise issues of evaluation during reference transactions.

Another example of how the results can improve our understanding of reference service was that the providers' perceptions were lower than those of the users by about one half point on the five-point scale. The differences were not immense, but the gaps were consistent and statistically significant. This could be a reflection of the different perspective of the learner and the teacher, or of either over-confidence on the part of the user and under-confidence (perhaps realism?) on the part of the provider, or both. Whatever the case, it provided a good opportunity to raise morale. Reference providers can be pretty hard on themselves. The nature of the job—dealing every day with struggling users and repeat questions—can give us the impression that we are making no progress. Showing how satisfied our users were and that we consistently rated ourselves lower was a positive lesson for our reference providers.

One final example from our test: the users' mean responses for transactions involving a librarian were consistently higher than for those involving staff. Again, from a management perspective, this is a good thing to know. Librarians are expensive. This result may help us justify why we have them on the reference desk instead of less formally qualified staff. Our users got more help and were more satisfied with the results in transactions involving librarians rather than staff.

## CONCLUSION

As noted earlier, this is a study of perceptions. The questionnaire cannot be used to objectively evaluate the help the user received or the quality of the teaching involved in the transaction using external criteria. The results of this survey in Pittsburgh academic libraries confirmed LibQUAL results—users generally give public-service personnel high marks for service. More specifically, this study indicates that our users perceive that they get the help they need and are satisfied with reference service. Users are more positive about their interactions with librarians than staff, though they are positive about both. But they are slightly less satisfied with our teaching and particularly evaluative assistance. Reference providers (in our test, librarians most of all) are harder on themselves and users (or perhaps they are more realistic) than users are. If administrators rely on reference providers' perceptions of service satisfaction and effectiveness, they will have a distorted view of the service. Gathering responses from both sides of the reference desk provides a more nuanced picture of reference service as a whole.

This method for evaluating reference could easily be extended to synchronous online reference transactions, delivered to user and provider at the end of the transaction, and even to e-mail reference, but it would be more intrusive and time consuming to expand it to phone reference. It has been tested in a variety of academic libraries, but there is no reason why it cannot be used by public libraries that follow an instructional model of reference service. The combination of perspectives from users and providers adds value to the instrument, but it is possible to use the user portion alone as an even simpler evaluation tool. Once developed, it was relatively easy to administer and did not unduly burden reference service providers or users. It was also relatively easy to process the data and analyze the results. If used to test online reference, data-processing time could be reduced even further with automatic collection of digital responses using readily available software and systems.

Like many other libraries, the University of Pittsburgh Libraries conducts the LibQUAL survey on an annual basis. More specific evaluations of reference service should be conducted on a similarly regular basis, if not continuously. As a manager, I would also like to be able to evaluate the service performance of individual reference providers. This would enable the annual or tenure reviews of reference service personnel to be based on something other than observation, anecdote, and the occasional thank you note from a user. Such individual evaluation would make reference service evaluation more comparable to the student evaluations that are a common component of classroom faculty review processes. Obviously there are delicate personnel issues involved in such a step. Also, individualizing the questionnaire might add to the problem of a bias towards positive transactions. But I think the librarian profession needs to move beyond anonymous evaluation and have the confidence to assess the service quality provided by individual librarians, particularly if we can find a way to do so quickly and easily.

## ACKNOWLEDGEMENTS

The author wishes to thank Jean Alexander, Head of Reference at the Hunt Library at Carnegie Mellon University; Jill Ausel, Director of the Library at Chatham College; and Elaine Rubinstein of the Office of Measurement and Evaluation of Teaching; Rachel Shackelford, Hillman Reference and Instruction Office Manager; Kate Thomes, Engineering Librarian; Eve Wider, Hillman Reference Area Supervisor; all of the University of Pittsburgh. This article is based on research conducted while the author was head of Hillman Public Services at the University of Pittsburgh and first presented at the Reference Research Forum at the ALA Annual Conference, June 25, 2006.

## References

1. Reference and User Services Association, “Definitions of a Reference Transaction,” www.ala.org/ala/rusa/rusaprotools/referenceguide/definitionsreference.cfm (accessed June 6, 2007).
2. Reference and User Services Association, “Guidelines for Behavioral Performance of Reference and Information Service Providers,” www.ala.org/ala/rusa/rusaprotools/referenceguide/guidelinesbehavioral.cfm (accessed June 6, 2007).
3. LibQUAL Charting Library Service Quality, “Welcome to LibQUAL,” www.libqual.org (accessed June 1, 2007).
4. Matthew L. Saxton and John V. Richardson Jr., *Understanding Reference Transactions: Transforming an Art into a Science* (New York: Academic Pr., 2002).
5. Jo Bell Whitlatch, *Evaluating Reference Services: A Practical Guide* (Chicago: ALA, 2000).
6. Jo Bell Whitlatch, *The Role of the Academic Reference Librarian* (New York: Greenwood, 1990).
7. Marjorie E. Murfin and Gary M. Gugelchuk, “Development and Testing of a Reference Prediction Assessment Instrument,” *College and Research Libraries* 48 (July 1987): 324–25.