

Quality Control: a Necessary Good for Improving Service

Firouzeh F. Logan

Krystal Lewis

Abstract

This paper reports on the benefits, challenges, and effective methods of assessing the quality of virtual references services at an academic library.

Keywords: Academic libraries, assessment, evaluation, virtual reference services

Firouzeh F. Logan
University of Illinois, Chicago
Flogan@uic.edu
Head of Reference, Assistant Professor
R M Daley Library
801 S Morgan
Chicago, IL 60607
312-996-2735

Krystal Lewis
Temple University
kmlewis1@temple.edu
Coordinator of Information Literacy and Reference
Samuel L. Paley Library (017-00)
Room: 122
1210 Polett Walk
Philadelphia, PA 19122
215-204-3187

Introduction

In 2006, when the University of Illinois at Chicago (UIC) had to renew its virtual reference management service contract with OCLC's Question Point, the UIC Library had two options: either join a global reference cooperative for a substantially reduced cost, or renew a much more expensive individual account. The administration decided that UIC's own librarians would best serve UIC's users, and the Library declined to join the cooperative. The discussion at the time was extensive and the decision unanimous, but it also raised a question that no one could answer with any certainty: was the Library's virtual reference service good enough for such a cooperative, and, more importantly, good enough for UIC's own users?

In these difficult budgetary times, it is more important than ever to justify the expense of every service. Assessment allows not only for informed decision-making about resource allocation, but also provides information about ways to improve service. It helps determine what percentage of a budget should be allocated to the actual virtual reference service and training as opposed to the cost of the virtual reference software and the electronic resources that are used to answer questions. More to the point, most librarians would argue that regardless of the economics, it is always good policy to determine whether a service is meeting its goals and objectives. In this digital age, clients have high expectations, and effective service gives library clients a positive experience of the role of the Library in their academic lives.

The Richard J. Daley Library's reference department, the largest department participating in the cooperative digital reference service at UIC, decided to design and implement an assessment project that would set service standards first, and then would

use those standards to improve the service and staff training. The project could then perhaps also serve as a model for the other eight departments that participated in the service. As daunting as the process seemed, we felt that these steps would insure the future viability of the service and would be well worth the time and effort.

History

The University Library at UIC is geographically dispersed and serves specific subject specialties. Four of the six libraries are Libraries of Health Science, one is a Science library, and the other, Daley Library, serves the liberal arts and sciences, the arts and architecture, and professional programs such as Business. From 1987 until 2003, each library's reference department had its own email address and answered its own questions. In 2003, this online reference service was changed from a decentralized service to a collaborative service with dedicated staffing, marketing and evaluation of the effectiveness of the shared service.

In order to serve all six library sites in a more efficient, flexible manner, the Library adopted OCLC's Question Point. An implementation task force set about examining email systems that met the requirements that were important to the Libraries (DeGroote *et al.*, 2005):

- Centralized system
- Seamless interface for the patron
- Allow subject specialists to answer questions no matter where they might be directed

At that time, 35 librarians participated in the shared service. The task force examined the effectiveness of changing from multiple service points to a single service point, and the members decided to measure the success of the latter approach by examining the data collected in each transaction: question type, user status, subject area, etc. The study concluded “a centralized digital reference service is feasible even in a complex academic environment serving a diverse user population with widely divergent disciplines spread across five campuses” (DeGroot *et al.*, 2005, 451). The implementation group was aware that transcripts would allow for evaluation of accuracy and completeness of answers. However, questions about the quality of the answers were not pursued at that time.

Literature Review

Even though virtual reference has been adopted at a stunning rate, assessment of digital reference is still in its infancy. Early on it became obvious that in the emerging digital environment, digital assistance would be a viable alternative to face-to-face and telephone service for providing quality reference service to users. According to Joseph Janes, in 2000, 45 percent of academic libraries offered some type of digital reference service (Janes, 2002). In 2004, according to the National Center for Educational Statistics, that number was 69 percent, and in 2008 it was 72 percent.

But as late as 2002, Janes noted that only 9 percent of survey respondents indicated that they did any kind of evaluation. Today, the library community recognizes that the ability to communicate with users in a variety of ways is very important: email, chat, text, phone, and in person. At times it might be important to know why people use the medium they choose, but as far as assessment goes, it is always important to know

whether high quality service is being delivered. Given the economic constraints facing academic institutions, it is more important than ever to examine every service the library offers.

In preparing for the virtual reference assessment project, the authors reviewed the literature and found that many professionals besides librarians need to demonstrate that they provide high quality service and therefore engage in peer assessment. This includes legal aid programs in the United Kingdom (Paterson, 2007), the medical profession and the boards of hospitals (Combes, 2009), and educators (Langan, *et al*, 2005). They all debate the meaning of quality and the criteria used to measure quality, but none provided a template for this project.

Just as there are well-documented issues with assessing face-to-face reference, a range of issues also faces assessing digital reference. Research on the kinds of questions that are best answered by digital reference has yielded no definitive findings about whether the type of library (special, academic, public, etc.) is significant when looking at quality, or whether librarians provide evidence of evaluation criteria and authority by identifying their resources. It is not clear if any of these distinctions makes a qualitative difference. There are studies looking at procedural issues, such as service accessibility, user awareness, response time, the times of use, user satisfaction, and cost effectiveness, but little scholarly attention focuses on the quality of response (accuracy, appropriateness, instructiveness, etc.): “There are no consistently held quality criteria in the library profession for reference over the Internet” (Lankes, McClure, Gross, 2003, 325).

Three valuable studies address the accuracy and completeness of answers. In a 2003 article, Marilyn White, Eileen Abels, and Neal Kaske reported on a pilot study in both academic and public libraries that sought to test various methodologies and approaches to evaluating chat reference. Some of the measures include: descriptive statistics, user satisfaction, peer review, the cost of service, and staff time expended. No single method eclipsed the others, but the study helped to inform UIC's selection of criteria to examine in its own project. The NCknows evaluation study published in 2005 advocated using peer reviews of chat transcripts (Pomerantz, *et al.*, 2006) to determine whether questions were answered completely and correctly. The Arnold and Kaske study also published in 2005, examined transcripts of chat sessions at the University of Maryland in College Park (Arnold, Kaske, 2005). In both studies, librarians worked independently studying and evaluating the answers in the reference transcripts.

Pomerantz's 2008 article addresses content analysis, and offers helpful advice about "rigorously defining what constitutes an accurate and complete answer or an instance of instruction or an expression of thanks" (Pomerantz, 2008, 16), but it neither says who should do the defining nor how the defining should be done. Shachaf and Horowitz discuss the effectiveness of digital reference, especially as measured by adherence to RUSA and IFLA guidelines, but they concentrate more on adherence to the guidelines than the actual quality of the responses.

David Ward's article, "Using virtual reference transcripts for staff training" provided both a few "how to's" and a few "how not's." For example, like Shachaf and Horowitz, he advised using RUSA guidelines to gather criteria for an evaluation checklist, but as the participants in his study were graduate students, and not tenure track

librarians, UIC decided on a more inclusive and consensual method that would result in a better understanding of and commitment to the project and its goals. One homegrown idea was to have the staff participate in the process of creating the evaluation model from the beginning, which encouraged buy-in to the project.

Methodology

In designing the assessment, we understood that there would be objections and impediments that needed to be met and overcome before the project could even begin. At first when it was announced that email reference service would be evaluated at UIC, the librarians resisted. They felt this was a particularly intrusive process, and no one wanted this very detailed appraisal to affect annual performance or tenure reviews. However as it became clear that the objective of the evaluation was neither to penalize nor to grade each librarian's performance, but rather to set quality standards, identify best practices, and improve service, resistance subsided, and everyone agreed that it would be more productive to formulate an approach that would yield valid and useful assessment data and would be acceptable to the librarians being evaluated.

Surprisingly, this first thorny issue was rather easily dealt with by reassuring everyone that the project could not affect annual performance reviews, as the emails would be stripped of all identifying material prior to being examined, and there would be total transparency in the process. That is, everyone would examine transcripts together at specially designated department meetings. Also reassuring was that one of the stated goals of the project was to develop standards and best practices. As everyone has an interest in good service, it was decided that the project would also provide an opportunity

to develop a training module for all the libraries, which would not only train new librarians, but also allow for periodic evaluation of the quality of email reference service.

The other rather formidable task was deciding which standards the Library should use to judge the quality of the answers. It was really difficult to obtain consensus on the standards that were most important and appropriate. In order to build consensus and legitimacy, we decided that the department would select the standards used to evaluate each answer as a group. There had to be agreement as to the definition of “quality.” In addition to the definition, there had to be consensus on what acceptable “quality” was, and what was not acceptable. Ultimately, the group concluded that the most important criterion should be whether or not user information needs were met, but it was also agreed that successful transactions were more than just mastering procedures, and that accuracy, tone, good grammar, authoritative sources, and a proper reference interview were also important components of an answer. Ultimately, academic librarians must be good guides to information sources.

Once initial objections were addressed, the Head of Reference began selecting and preparing transcripts for analysis. (Appendix 1). She randomly selected 10 percent of each reference librarian’s email and chat transcripts from a three-month period. The Head chose a large enough sample of transcripts to represent adequately the diversity of questions fielded by the reference librarians in the department. She then downloaded the transcripts into text files, where she stripped them of any indentifying patron information such as name, email, and university status. She also removed any information that could be used to identify the librarian answering the query, such as signature files and

references to subject specialization. Over the course of three meetings, the Reference department analyzed and discussed 30 email and chat transactions.

Data and Analysis

The first step in assessing the transactions was to determine the type of question being asked. The categorizations included:

- Directional – questions that require assistance with locations, schedules, etc, but no knowledge of information sources
- Ready reference – questions with easily accessible answers
- Research – questions that involve the knowledge of information sources and require a rather high level of complexity
- Instructional – the question requires instruction in the use of a resource or conducting research, or the librarian provided instruction
- Technical/Access – questions about access to resources (off-campus/proxy, subscriptions, outages)
- Circulation/ILL/Reserves – questions related to borrowing materials
- Other – Any questions that don't fit into the above categories; mostly comments or suggestions (Table 1).

Table 1: Types of Questions*

Question	Number	Percentage
Directional	5	12%
Ready Reference	13	31%
Research	10	24%
Instructional	8	19%

Technical/Access	3	7%
Circulation/ILL/Reserves	3	7%
Other	0	0%
Total	42	100%

*Though 30 transcripts were analyzed, some contained more than one question, so the total number of questions is 42.

Determining the type of question asked seems like a fairly straightforward task, but it proved otherwise. This was at first surprising, but less so as other studies have noted similar difficulties (Murphy, Moeller, Page, Cerqua, & Boarman, 2009; Neville & Henry, 2009). Often the reference librarians disagreed, and there was dispute over whether a question was directional or a technical/access question. Librarians debated about the level of complexity necessary for a question to be considered a “research” question and whether a user was asking for or needed instruction. While reviewing a transcript and taking a user’s question at “face value,” it is easy to see how it might be difficult to be certain about the question type. When a person asks a question at the reference desk, a librarian would presumably conduct a reference interview to determine the person’s information need. In most of the transcripts examined, a reference interview was not conducted and the librarian replied with an answer.

The lack of reference interviews in virtual, especially email, reference might be due to the responding librarian’s desire to give the user an adequate answer as quickly as possible. Indeed, conducting a complete reference interview via email can be time-consuming and could easily take several days. However, the lack of a reference interview meant there was little contextual information to help when determining the type

of question. This accentuated the importance of conducting a reference interview, even in virtual reference. If a group of reference librarians cannot even agree on the type of question a user is asking, it is possible then that a librarian might misinterpret the question itself and not provide the needed answer. And if the user is not receiving appropriate answers, then the quality of the virtual reference service itself is in doubt. Good communication skills are key. The clarification a reference interview can provide might be essential to a job well done.

To assess the quality of the responses, the department considered the RUSA Guidelines for Behavioral Performance of Reference and Information Service Providers (<http://www.ala.org/ala/mgrps/divs/rusa/resources/guidelines/index.cfm>). From these guidelines, the department developed questions to analyze and discuss each transaction:

- Is the response correct and complete?
- Are the tone, grammar, and personal style appropriate?
- Is the answer instructive and is that appropriate?
- Is there any follow up and is that appropriate?

These questions framed the discussion and determined if and how a response met each specific quality or criterion (Table 2).

Table 2: Quality of Responses

Quality/Criterion	Number of acceptable responses	Percentage
Correct and Complete Response*	22	73%
Correct Grammar	29	97%
Courteous in tone	25	83%

Instruction Provided	18	60%
Referred	4	13%

*Response met both criteria; responses that met only one of the two were not counted

The discussion about quality alone made it clear that there are different levels of acceptable quality. Factors outside the guidelines selected such as the patron's affiliation and the patron's status did impact the answers and added to the complexity of trying to match the quality of service to the goals and objectives of the department. For example, some felt that the librarian should not be obliged to conduct as thorough a reference interview or use the opportunity to instruct if the person were not affiliated with the university.

Although identifying the type of question asked caused some concern about the quality of the virtual reference service, further analysis allayed those concerns. Librarians almost always used proper grammar (97 percent) and they usually struck a courteous tone in their answers (83 percent). Librarians provided instruction to people in 60 percent of their responses and answered questions both correctly and completely 73 percent of the time. Reference librarians were usually able to answer patron inquiries themselves, referring only 13 percent of the questions to other librarians or departments. The quality assessment actually made the department feel confident that the virtual reference service provided was quite good.

Unexpected Benefits of the Project

The examination of the quality of the transcripts produced a few important side benefits. One was a renewed appreciation for maintaining the privacy of patrons and the confidentiality of their inquiries. Concern for maintaining their own anonymity seemed

to sensitize the librarians to the significance of not compromising a patron's right to privacy and confidentiality. Another unexpected benefit was an increased appreciation for the resources available at UIC. While considering the accuracy of a response, librarians' awareness of the sources used and those that could potentially be used was raised. This was especially helpful as the department included librarians with a wide range of experience and knowledge. Librarians discussed the merits of both newer and older sources and shared unknown sources with each other.

Not only did the librarians share their favorite sources, but over the course of the project they also disclosed their identities and personal experiences. As previously stated, in order to facilitate a freer discussion, steps had been taken to protect the identity of the answering librarian. But as the department considered each transaction, librarians took ownership of their work and revealed their identity. They were keen to provide further context in which to judge their work, e.g., why a particular source was chosen to answer the question, or why or how follow-up had occurred. Just as the librarians in the department may have missed the nuances in patrons' questions by taking the transcripts at "face-value," so, too, they sometimes failed to see the rationale behind the manner in which librarians answered the questions. Sometimes librarians provided follow-up to an answer in person or over the phone, and that was not reflected in the transcript. Although virtual reference transcripts do offer a unique opportunity to assess the service, they have this and other limitations.

By participating in these group evaluations of their reference transactions, the librarians in the project became better acquainted with the library's resources and with

each other and appreciated more fully the breadth and complexity of their work as librarians.

Next Steps

After the successful conclusion of the project, the University Libraries' Public Services Advisory Council, PSAC, decided that this evaluation should be undertaken at the department head level throughout the University Library's departments and sites. PSAC used the Daley Library checklist as a starting point for building a new evaluation checklist. Today the new checklist is used in training all new public service librarians and not only provides them with the Library's standards but also spells out the Library's expectations to these new service providers (Appendix 2).

Thanks to the work done at the Daley Library, setting the standards proved to be the easy part of this next phase. It was understood that it was in everyone's interest to take the time to examine the service that each department rendered (Appendix 3).

The goal of having a system in place for training both service providers and service evaluators was met, and more importantly the department met a very crucial RUSA guideline: "Facilitate regular assessment of the virtual reference program's effectiveness by library staff and administration" – a most necessary goal in this environment of budgetary pressure and accountability.

Conclusion

Digital reference has become such an accepted, integral part of the reference services of the UIC Library that it is important to continue monitoring the service in order to maintain a high quality of service. The service has become incorporated into the information lives of both reference librarians and library patrons. No one knows or has

recorded when and what the first digital reference transaction may have been, but it is acknowledged that virtual reference service is a vital part of the public services of every university library with a dedicated staff, training, marketing, and assessment of sorts.

The inclusive process UIC used to design and implement the assessment project not only helped tremendously in eliminating suspicions that the evaluations would be used unfairly, but also helped in building consensus and in enlisting the necessary cooperation. It was important that the very librarians who offer the service work together as a group to shape the program and decide what constitutes a successful transaction. The process built awareness of the factors that make for positive online transactions. It quickly became apparent that it is essential to set quality standards in order to measure performance and identify areas for improvement, and that it is important to educate old and new staff about expected standards of service. The study found that the department was meeting its goal to provide high quality, high impact service that meets user needs. Having solid data demonstrating its strong service proved to be a huge boost to departmental morale. The project has proved to be of an even more universal benefit by creating a template for use by other departments and sites who are interested in identifying strengths and weaknesses in their existing service.

Appendix 1: Process of Transcript Review and Evaluation

1. Select Transcripts

- Randomly select a percentage of each person's transcripts and logs
- Select a large enough sample to adequately represent the diverse questions we are asked

Person responsible: department head or a committee/task force.

2. Prepare Transcripts

- Strip patron identifying information
- Remember to strip librarian identity as well – this allows for a much freer discussion
- Must stress that individuals are not being evaluated; the work is

Person responsible: department head or a committee/task force.

3. Set Standards

- Quality standards should be decided upon and acceptable norms established
- It is most useful to come to an agreement as to what constitutes “quality” reference service for several reasons:

1. Establishes a standard for future training

2. Makes clear that there are different levels of acceptable quality
 3. Underlines the complexity of matching the quality of service to the goals and objectives of the department
- Select standards for judging from the following RUSA standards:
 1. Courtesy/tone
 2. Accuracy
 3. Interview/clarification
 4. Information literacy
 5. Response time
 6. Grammar/spelling
 7. Follow up/referral
 8. Patron satisfaction

Person responsible: entire department.

4. Evaluate

- Identify type of question
 1. Ready reference
 2. Research
 3. Technical / Access
 4. Circulation/ILL/Reserves
 5. Directional
 6. Instructional

7. Other

- Analyze and discuss – suggested questions
 1. Is the response accurate?
 2. Are the suggested resources valid?
 3. Are the tone, grammar, and personal style appropriate?
 4. Is the answer instructive and is that appropriate?
 5. Is there any follow up and is that appropriate?
 6. Is the answer complete?
 7. Are standards met?

Person responsible: entire department.

5. Follow-up

- Identify areas for improvement
- Modify e-mail reference training procedures

Person responsible: department head or a committee/task force.

Appendix 2

Department Head Virtual Reference Training

Ask A Librarian

Evaluation Checklist

Requisite Quality Measures	YES	NO	N/A
Timeliness of Response			
Salutation/Signature/Contact Information			
Courtesy/Tone			
Grammar/Spelling			
Accuracy of Response			
Completeness of Response			
Appropriate resources selected			
Fair use copyright practices followed			
Question within librarian's expertise			
Optional Quality Measures			
Instruction provided			
Answer appropriate to user audience			

Answer in easy to read blocks of text or outline format			
Defined terms and avoided jargon			
Links to online resources supplied			
Sources fully cited			
Database search citations provided			

Appendix 3

Part 2 – Evaluation Training

1. Guidelines
 - A. Cover guidelines
 - B. Handout
 - C. Checklist
2. Three emails stripped and ready for evaluation
 - A. Late answer
 - B. Salutation/Closing remarks missing
 - C. Incorrect answer
3. Search
 - A. Select from Closed and Service Record
 - B. Search by number
 - C. Search by librarian
4. Final Comments
 - A. Use own good judgment – but evaluate regularly
 - B. Training an issue – speak to PSAC
 - C. If answer is not correct, be sure someone follows up with patron

References

- Arnold, Julie and Neal Kaske. 2005. Evaluating the Quality of a Chat Service. *portal: Libraries & the Academy* 5: 177-193.
- Combes, John R. 2009. Peer Perspective Deepens. *H&HN: Hospitals & Health Networks* 83: 56-56.
- De Groote, Sandra L., Josephine L. Dorsch, and Scott Collard. 2005. Quantifying Cooperation: Collaborative Digital Reference Service in the Large Academic Library. *College & Research Libraries* 66: 436-454.
- Janes, Joseph. 2002. Digital Reference: Reference Librarians' Experiences and Attitudes. *Journal of the American Society for Information Science & Technology* 53: 549-566.
- Janes, Joseph. 2008. An Informal History (and Possible Future) of Digital Reference. *Bulletin of the American Society for Information Science & Technology* 34: 8-10.
- Langan, A. M., C. P. Wheeler, Emma M. Shaw, Ben J. Haines, W. R. Cullen, Jennefer C. Boyle, David Penney, et al. 2005. Peer Assessment of Oral Presentations: Effects of Student Gender, University Affiliation and Participation in the Development of Assessment Criteria. *Assessment & Evaluation in Higher Education* 30: 21-34.
- Lankes, R. D., Melissa Gross, and Charles R. McClure. 2003. Cost, Statistics, Measures, and Standards for Digital Reference Services: A Preliminary View. *Library Trends* 51: 401.
- Murphy, Sarah Anne, Sherry Engle Moeller, Jessica R. Page, Judith Cerqua, and Mark Boorman. 2009. "Leveraging Measurement System Analysis (MSA) to Improve Library Assessment: The Attribute Gage R&R." *College & Research Libraries* 70, no. 6: 568-277.
- Neville, Tina M., and Deborah B. Henry. 2009. "Reference Classification--Is It Time to Make Some Changes?." *Reference & User Services Quarterly* 48, no. 4: 372-383.
- Paterson, Alan. 2007. Peer Review and Quality Assurance. *Clinical Law Review* 13: 757-778.
- Pomerantz, Jeffrey. 2007. Evaluation of Online Reference Services. *Bulletin of the American Society for Information Science and Technology* 34:15-19.
- Pomerantz, Jeffrey, Lili Luo, and Charles R. McClure. 2006. Peer Review of Chat Reference Transcripts: Approaches and Strategies. *Library & Information Science Research* 28: 24-48.
- Schachaf, Pnina and Sarah M. Horowitz. 2008. Virtual Reference Service Evaluation: Adherence to RUSA Behavioral Guidelines and IFLA Digital Reference Guidelines. *Library & Information Science Research* 30: 122-137.
- White, Marilyn Domas, Eileen G. Abels, and Neal Kaske. 2003. Evaluation of Chat Reference Service Quality. *D-Lib Magazine* 9.