EVALUATION AND ANALYSIS

Evaluating Electronic Reference Services:
Issues, Approaches and Criteria

Eric Novotny

SUMMARY. As electronic reference services become routine in many libraries, it is time to systematically examine how they are being implemented and used. Unfortunately, few libraries have rigorously examined their electronic reference services. We still know very little about who uses electronic reference services or why. We also do not know how satisfied our users are with the new services we are providing. This article provides an overview of the chief methodologies available for conducting assessments of electronic services (e.g., surveys, usability studies, observation, etc.). Existing criteria for evaluating reference services are discussed with suggestions for how they can be applied or adapted to the online service environment. *[Article copies available for a fee from The Haworth Document Delivery Service: 1-800-342-9678. E-mail address: <getinfo@haworthpressinc.com> Website: <http://www.HaworthPress.com> © 2001 by The Haworth Press, Inc. All rights reserved.]*

Eric Novotny is Humanities Librarian, Arts and Humanities Library, University Libraries, The Pennsylvania State University, University Park, PA 16802-1803 (E-mail: ecn1@psu.edu).

INTRODUCTION

While still a relatively recent development, the provision of reference service via electronic means is by now well established in the majority of American libraries. Ninety-six percent of the libraries responding to an Association of Research Libraries survey reported that they offered reference assistance electronically.[1] For the purpose of the survey, electronic reference was defined as "reference service designed for remote users and identified by a specific link from a library's website." Thirty percent of these services were more than five years old, while another fifty percent had been established more than three years ago. Only a small minority of libraries had established their online reference services in the last year.[2]

As electronic reference matures it is time to take a more pragmatic look at the way it is used. Unfortunately, very few libraries have consistently and rigorously examined their electronic reference services. Only 13% of the libraries responding to the ARL survey indicated that they conducted any assessment of their electronic reference services. Judging by the sparse professional literature on the topic, very few of the assessments that have been done have made it into print.[3] This has left librarians and administrators with more questions than answers. We still know very little about who uses electronic reference services or why they choose to do so. We also do not know how satisfied our users are with the new services we are providing.

The goals of this paper are: (1) to convince librarians that they need to conduct ongoing assessment of reference services, including those delivered online, (2) to provide an overview of the chief methodologies available for conducting such assessments, and (3) to review existing criteria for evaluating reference services and suggest how they can be used to judge the efficiency and effectiveness of electronic reference services. It is hoped that readers will not only be better informed, but also inspired to closely examine their services and ways to improve upon them.

Before proceeding it is prudent to clarify terms. What is a reference service generally, and more specifically, what is an electronic reference service? There are many definitions for both terms in the library literature. Some emphasize the act of answering questions, while others are broader. The more expansive definitions include pathfinders, bibliographies, FAQ databases or tutorials as reference services. For this article I will rely on a definition formulated by Chris Ferguson and Charles Bunge. They identify the defining characteristic

of "traditional" reference service as "answering questions posed by users."[4] This model emphasizes human attention and interaction. It excludes non-human, purely technological approaches to public service. Resources such as databases and instruction modules also fall outside the scope of this article. I believe this definition of reference service captures the essence of the reference encounter, whether in-person or online. The interaction between users and information professionals remains the most important relationship, and the one most deserving of attention. Of course in order to be an electronic reference service the information must be conveyed electronically in some way, either via e-mail, chat room discussion, web form, or some other from of digital communication.

Evaluation also requires a precise definition. Much like "electronic reference services" the term can be used to describe many things. Evaluation can include everything from a telephone survey, to gathering usage statistics, to making a mental note of a user's comments at the reference desk. Charles McClure defines evaluation as "the process of identifying and collecting data about specific services or activities, establishing criteria by which their success can be assessed, and determining both the quality of the service or activity and the degree to which the service or activity accomplishes stated goals and objectives."[5]

WHY EVALUATE?

The evaluation process described by McClure is a time-consuming one. It is not enough to simply conduct one study and analyze the results. Evaluation must be integrated into the library's operations, and built into the implementation of any new service. It is also not enough to merely gather quantitative measures such as the tabulation of reference desk statistics. To truly evaluate a service it is necessary to gauge the quality of what is being offered as well as the quantity. Quantitative measures tell us little about the value of the service to the users. An increase in recorded desk activity may not be a positive thing. It may reflect greater user frustration with a poorly designed library.

Quality can be a difficult concept to measure, requiring greater efforts on the part of investigators. Why do we need to expend this effort to evaluate electronic reference services? After all, isn't it obvious that high-tech service provision is the wave of the future? This may well be, but it does not release us from our professional obligation to provide the highest level of service to our users. New services do not magically appear without cost. Every new offering represents a loss of resources in some other area. An hour spent staffing an online chat service is an hour that cannot be used to accomplish other important library goals.

Any introduction of new services should be done with a clear accounting of the costs and benefits involved. To paraphrase a famous saying, there is no such thing as a free service. Even if there is no money expended, the involvement of staff represents an opportunity cost to the institution. The opportunity cost is the value of the services that could have been provided with the time and resources devoted to the new service.

Some librarians may counter that they are not cutting other services, that electronic reference services are simply being added to their existing workloads. If true, this is not without its own costs. The phenomenon of librarian burnout is hardly a new one, but technology seems to have exacerbated the problem. The trend towards increasing workloads and expectations was identified as early as 1984 by Bill Miller. In his article entitled "What's Wrong with Reference: Coping with Success and Failure at the Reference Desk," Miller noted that librarians are expected to maintain existing services while adding new ones. In an effort to keep pace with rapid technological changes, librarians found themselves working longer hours. This was not a sustainable solution. The end result was an increase in stress and a decline in job satisfaction.[6]

Ironically the problem of burnout is caused by our successes, not our failings. Librarians have emerged as leaders and experts in the information realm. We have eagerly taken on new responsibilities, while at the same time maintaining our existing services. This is only natural. It is often easiest to delay hard decisions by offering a new service without eliminating or reducing existing responsibilities. While this may work for a while, it is not a recipe for long-term job satisfaction. It is not possible to work longer and harder indefinitely.

Inevitably choices need to be made. We can offer many services that are of value to our users. We cannot, however, provide every potentially useful service our users might want. In an environment of fixed resources, libraries cannot be all things to all people. There is a need to prioritize amongst the many service options available. A consistent, coherent program of evaluation can make choosing amongst the available options easier by providing sound data on which to base decisions.

EVALUATION METHODS

While many librarians acknowledge the benefits of evaluating their services, they face obstacles that prevent them from implementing such a program at their institutions. Designing and carrying out an evaluation project can be a daunting task. Fortunately a number of guides exist which offer advice on how to get started. The *Reference Assessment Manual,*[7] compiled by the ALA Evaluation of Reference and Adult Services Committee, examines at the state

of assessment in all areas of reference services. This includes training, the reference environment, what reference services should be offered, costs and outcomes, and reference effectiveness. A more recent text that updates the *Reference Assessment Manual* is *Evaluating Reference Services: A Practical Guide* by Jo Bell Whitlatch.[8] As the title suggests, this book is aimed at busy librarians in the field. It offers advice on how to plan an evaluation, select an appropriate research tool, and deliver quality reference service.

It is not possible for this article to describe in detail the many research methodologies available for evaluating reference services. It is desirable, however, to briefly present the advantages and disadvantages of each approach. The most widely used research methodologies are: case studies, focus groups, individual interviews, and surveys. Other methods that have been used to assess electronic reference services are cost-benefit analysis, usability studies, web log analysis, and statistics. The most appropriate method for your situation depends on the type of information you are seeking. Do you want a large number of responses, or would you prefer more in-depth analysis from a few sources? Do you want emotions and opinions, or would you prefer to collect numbers that are easily tabulated and analyzed? These are the sorts of questions you need to consider as you ponder which methodology to use.

- *Case Studies*–A case study is an intensive examination of a particular situation or environment. The intent is to gain a deep understanding of the individual circumstances in a unique setting. A well-done case study employs a variety of methods, e.g., surveys, interviews, observations to thoroughly explore the issue being examined. The main drawback is that the narrow focus may mean that the results are only relevant to the particular case being researched.[9]
- *Cost-Benefit Analysis*–This type of analysis seeks to measure the value of services, or the cost of reducing an existing service. Various methods have been devised to estimate user benefits. One approach is to ask patrons to place a dollar value on a service such as personal assistance at the reference desk. Other measures may be substituted for dollars, such as asking users how a change in service might affect their happiness, efficiency, or productivity. A related methodology is contingent valuation method in which users are asked about their willingness to pay for library services. Conjoint analysis involves asking users to rank competing service levels or options. The advantage of these methods is it shows what users value most. It also assists librarians who must make choices about expanding or reducing programs, all of which have some merit. The disadvantage is that users may be overwhelmed by the complexity of the task. Considering and comparing more than a few programs at the same time can be mentally taxing for the average library user.[10]

- *Focus Groups*–A focus group is a group interview. Focus groups have the advantage of allowing users to interact, remind, reinforce, and reflect upon the views of others in the group. You can usually obtain richer, more detailed information than from a written survey. The disadvantages are that some individuals may be reticent to express their true opinions in front of others. Especially vocal members may hijack the session, dominating discussion. Avoiding these problems requires a carefully trained interviewer. Finally, it may be difficult to convince administrators and colleagues to act based solely on information obtained from such a small number of people. For these reasons focus groups are often used early in the process, and then supplemented with another research method such as a general survey.[11]
- *Individual Interviews*–Interviews have similar advantages to focus groups. Interviews allow participants to express their opinions and concerns in their own words. They allow you to explore a particular issue, i.e., why people do not use e-mail reference service, in considerable detail. Individual interviews do have their disadvantages. They can be time-consuming to set up and conduct. They also require careful staff training. A poorly prepared interviewer can seriously bias the results obtained. In particular, if you use staff members, they must be wary of defending existing policies or services from criticism. They must appear receptive to negative as well as positive comments.
- *Observation*–As the name suggests, this brand of research involves observing users in real-life situations. This may be done obtrusively, where the person being studied knows they are being observed, or it may be done unobtrusively without the knowledge of the other participants. The advantage of unobtrusive observation is that you are able to measure actual behavior, unlike surveys and interviews which involve people trying to predict their behavior or attitudes. The predictions may not always be accurate indicators of what someone will actually do in a given situation. A concern with the observational method is the potential for bias. Observer bias can taint the results, especially if staff enter into the study with preconceived notions of how users will behave. In an obtrusive observation, subjects who are aware they are being monitored may alter their behavior as a result of being under scrutiny. Finally, observational studies tell us what a person has done. They do not often reveal why they acted as they did, how they felt about it, or how they might act if presented with alternatives.[12]
- *Surveys*–Surveys are probably the most common data collection method used in libraries. Examples of surveys include a questionnaire distributed after an in-library instruction session, a user satisfaction form given to every person entering the library, a telephone survey of a random sample of the city's residents, and an e-mail feedback form on the library's web

page. The fixed format of most surveys provides consistency in responses. Everyone receives the same questions and the same set of instructions. The choices of answers can be standardized, making it possible to analyze a large number of responses efficiently. The anonymity of most survey methods can encourage honest responses. Surveys will not typically reveal the depth of information you can obtain from interviews or focus groups. Standardized responses may leave little room for respondents to accurately express their thoughts. Finally, question design is important. Respondents will usually not be able to ask for clarification if a question is poorly worded or ambiguous.[13]

- *Usability Studies*–Usability studies combine the features of observation and individual interviews. Usability studies have been conducted by libraries seeking to evaluate the design and utility of their web pages. Typically they involve giving a user a set of tasks, such as locating a book in the online catalog, or identifying where to go for assistance. Users are monitored while they attempt to complete the tasks. The session may be followed by individual or group interviews where participants are asked to explain their actions in more detail. The prominence and ease of use of an online reference service can be assessed using a usability study. The disadvantages to this approach are similar to those of focus groups. It can be difficult to set up a useful number of user sessions. Recording and analyzing the sessions can also be time-consuming.[14]

- *Web Log Analysis/Statistics*–Web log analysis is a variant of an observational study. Usage of a library's web page can be tracked. Basic data recording the number of times a page has been accessed, from where, and for how long can be relatively easy to obtain. It may be useful to determine which pages on your library's site receive the most traffic, and include links to your electronic reference services from these high-use pages. The downside of web logs and usage statistics is that they tell us little about the motivations of users. Web statistics can be inaccurate due to browser caching features that store frequently accessed pages on the user's own computer. Future visits to the same web page will not be counted if the page is obtained from the user's cache. If logs or statistics are used, they are best supplemented with surveys, or interviews.

Selecting a research tool is an important step, but before you can make a final decision, you will need to consider what exactly you are trying to measure. Are you trying to gauge usage, user satisfaction, librarian workload, or the efficiency of your service? This will inform your choice of methodology. You will also need to decide whether to conduct a quantitative study, or whether your focus will be more on the quality of the service.

QUANTITATIVE MEASURES AND BENCHMARKS

Quantitative measures are probably the most widely used approach to assessing electronic reference services. They measure how much, by whom, when, and where a service is used. Anyone who has made hash marks counting the number of questions at a reference desk is familiar with this form of assessment. Also well known are the drawbacks of this method. Knowing how many people asked a reference question, whether at a reference desk, or online, does not tell us much about the quality of the service. In fact, as has been noted earlier, more activity may be an indication that there is a problem in the library.

Statistics are also limited without a context to place them in. Are four e-mail questions a day indicative of a good, poor, or average service? How many questions does it take to make an electronic reference service a success? The answer to that question depends upon the goals and objectives of your service. If you are planning to conduct quantitative assessment, there are some recently published measures that can be used as benchmarks to compare your service.

Beth A. Garney and Ronald R. Powell conducted a survey of public libraries in the United States. Twenty-two libraries provided information about their e-mail services. The average number of e-mail reference questions received was 5.6. The modal, or most frequent response was 3 questions per week.[15] Academic libraries do not appear to be generating much more in the way of e-reference activity. A survey of Association of Research Library members revealed that the average number of questions submitted via formal electronic reference services was sixty-seven per month, or barely more than two per day.[16]

While libraries anticipate increasing demand for electronic reference services, the current numbers reported are almost uniformly low. An active electronic reference service would appear to be one that receives any more than a handful of questions per day. There are a few notable exceptions, however, that illustrate the potential audience for online reference services. The Library of Congress receives over 2,500 questions per month, surpassed in the library world only by the 3,200 questions per month sent to the National Library of Medicine.[17] Even this impressive volume of activity is exceeded by certain non-library based services such as AskERIC, which receives about 4,000 questions per month.[18]

The reasons some services are well-trafficked, while others receive minimal usage, deserves additional scrutiny from librarians. Many of the users of national services such as AskERIC have access to excellent local public or academic libraries. Judging by the relative levels of activity, libraries are not the first choice for many of their online patrons. Any evaluation of digital reference service should include some sampling of non-users to determine why it is that they are opting to bypass their local libraries and using other services in

such large numbers. Is the library service not advertised enough, is the link to reference assistance too hard to locate, are libraries seen as "uncool" by our Internet savvy users? Until we actually ask someone these questions, the answers will remain merely speculative. Libraries need more research to uncover the reasons why some services thrive while others are ignored.

QUALITY MEASURES

Of course statistics are not the only, or even the best, way to measure the utility of a library's services. To adequately assess most library endeavors it is necessary to go beyond the numbers. An effort must be made to determine the quality of the service as well as the quantity. Quality can be difficult to define and measure, but this has not stopped many librarians from offering their own suggestions. A somewhat utopian vision of electronic reference services was envisioned by Chris Ferguson and Charles Bunge. They urge libraries to move towards "delivering high quality reference through the network to all users at all times and from all locations."[19] Others have proposed more modest goals. Acknowledging the widespread phenomenon of limited usage of electronic reference, proponents insist that such services are a success because they extend the reach of the library. The Santa Monica Public expresses this sentiment when it asserts that ". . . anytime we can extend the reach of our services to our clients . . . we think it is a success."[20]

Most librarians are probably between these two extremes. Few would regard a service that does not reach "all users at all times and from all locations" as a failure. The low usage of our electronic services does concern many, however. With limited staff and personnel resources it is not prudent to consider a service a success merely because it reaches a few users. We need a better gauge of the costs and benefits of electronic reference services in order to efficiently allocate resources to maximize benefits to our users. For this we need to go beyond merely gathering statistics on the usage of electronic reference services.

Fortunately there are a number of criteria that can be used to assess new and longstanding services (see Table 1). Many of the methods developed for traditional reference services will still apply to evaluating newer offerings. A starting point for discussing what constitutes quality electronic reference services may be to use the values proposed by Ferguson and Bunge. They identify the features of traditional reference service quality as access, equity, and professional personalized service.[21]

How might these traditional service values be measured in the digital environment? For electronic reference services access can mean many things. The option of e-mail or chat reference service extends access to users who may not have approached a librarian previously. A recent study of e-mail reference

TABLE 1. Criteria for Evaluating Reference Services

FERGUSON AND BUNGE	RUSA ELECTRONIC INFORMATION GUIDELINES	RUSA BEHAVIORAL GUIDELINES	MCCLURE AND LAPOTA	SERVQUAL
Access	Response time	Approachability	Extensiveness	Reliability
Equity	Accessibility/Convenience	Interest	Efficiency	Assurance
Personalized Service	Value	Listening/Inquiring	Effectiveness	Empathy
	Effectiveness in meeting needs	Involves patron in search process	Service Quality	Responsiveness
	Effectiveness in anticipating needs	Follows-up	Impact	Tangibles
			Usefulness	

found that international students were heavy users of electronic reference services. These students may have felt more comfortable writing than speaking English.[22] It is speculated that the Internet can be used to reach other library users who may be reluctant to ask for help at a traditional reference desk. These groups might include the elderly, handicapped, or particularly shy individuals.[23] These examples illustrate the potential for reaching a larger group of users. Further research is needed to see if this potential is being reached, and if not, how it can be. One approach might be to target specific populations using surveys to ascertain users' attitudes toward the service.

Equity is the ethic of ensuring that all users are able to locate the information they need. This does not mean that everyone receives the same *amount* of service. The needs of a faculty member will often differ from those of an undergraduate. What is important is that each receives the level of service appropriate to their needs. The existing evidence suggests that electronic reference services may not fare well in evaluations based on considerations of equity. This inequity is due to socioeconomic factors that are largely beyond the control of libraries. While computer ownership and Internet access are increasing at a rapid rate, troublesome gaps persist. This is reflected in the user demographics of a recent survey of public libraries. The users of the e-mail reference service were primarily highly educated computer scientists and information professionals. It is unlikely that these groups have more of an information need than other segments of the population. More likely they have better access to computer technology. These results suggest that online reference services may not be reaching users who are in need of information services.

Measuring the equity of your service probably requires a combination of approaches. An analysis of your current users can be conducted if you collect demographic data from those submitting requests. This can be done easily if users are asked to complete a form as part of the service. From this you can determine if your questions are coming from a particular group of users. A focus group of selected non-users can help to determine if there are obstacles to use in the community. Libraries can engage in outreach efforts to members of the community whose information needs are not being met through the library's online information services.

Equity can also refer to the level of service that is offered. Can users expect consistent, high-quality assistance from your digital reference service? The equity of your digital reference service can be evaluated by comparing it to your reference desk service. Do you treat users of your online service in the same manner as you do your in-person users? Many institutions appear to have intentionally created distinctions between the service they offer at the desk, and the services they offer online. Forty-one percent of the ARL libraries surveyed indicated that they only accept basic, factual questions online. It is common for academic institutions to only accept questions from the general public if they pertain to local resources.[24]

It may be necessary to limit service for pragmatic reasons, but at a cost in quality. The ideal service environment is one in which a user of your electronic reference service receives the same consideration as someone who asks a question at the reference desk. When establishing policies the potential impact on the equity and quality of your service should be considered.

Other measures of reference quality focus on inputs and outcomes. McClure and Lopata suggest concentrating on the following areas:[25]

Extensiveness: How much of the service is provided? Measures of this would be the number of users and the type and status of users.

Efficiency: How well are the resources devoted to the service being used? How much does it cost to provide the service? One measure of efficiency would be to determine how much each online reference transaction costs. This could be compared to the costs of other alternative service options.

Service Quality: How good is the service provided? User satisfaction is one measure of service quality. Peer comparison is another method of determining if a service is of high quality.

Impact: Does the service make a difference to the user? For example, does the user produce a better paper as a result of the citations suggested

by the librarian? Users can provide an estimate of the impact of a service, or it can be observed through other means (improved grades, citation analysis of papers, etc.).

Usefulness: How well does the service meet the needs of the user? This information can be readily obtained using interviews or surveys to discover how well existing services are of value to library patrons. Conjoint or trade-off analysis can also be used to measure the relative usefulness of competing service delivery methods.

BEHAVIORAL ASPECTS OF QUALITY: THE LIBRARIAN-PATRON INTERACTION

A high quality electronic reference service requires a human touch. Automated systems and online help screens have their uses, but they can only go so far in assisting our users. This is not an opinion that is universally shared. Some have argued that in the digital age personalized assistance will be judged too costly and inefficient. Individualized attention from a librarian has been compared to handcrafted furniture, which was largely displaced when mechanized production of furniture became possible. The author of this analogy suggests that librarians face a similar predicament. We may find ourselves replaced by automation, not because computers are better searchers, but because they are faster, and cheaper.[26]

While automation may be inevitable, it appears that for the present most people still prefer human interaction and guidance. A recent marketing survey determined that 90% of online shoppers want human assistance available while they make their purchases.[27] If you want to measure the human element of your online service, there are criteria you can use. Some have been created specifically for the digital environment, while others need adaptation. Among the latter category are the RUSA Behavioral Guidelines (see Table 1). These were developed to describe the best practices for face-to face interaction at a reference desk. In many ways, however, they are applicable to an asynchronous reference environment.

The main evaluation criteria for the RUSA Behavioral Guidelines are Approachability, Interest, Listening/Inquiring, Involving the patron in the search process, and Engaging in follow-up. Approachability in an online environment translates to how easily your service is located, and how user-friendly it is. A long, complicated form can discourage patrons from submitting their questions. The appearance of the page can suggest approachability, or it can act as a deterrent. Many pages devote their first paragraphs to telling users what questions cannot be asked, and who cannot use the service. This is the electronic equivalent to sitting at the reference desk with arms folded and a

stern look on your face. The location of the page can also mimic approachability. Is your online service as easily located as your reference desk? Most users can probably identify your reference desk, but are the online service links as obvious?

Interest can be harder to determine. In an online environment the closest characteristic may be response time. Users do not necessarily expect an instant response, but a rapid response is a way to assure the patron that the question is important to the library. This does not mean that the question must be *completed* quickly, simply that initial contact be made in a timely fashion.

Listening and inquiring relate to the reference interview. It is well established that the reference interview can be difficult to conduct online. What takes place instantly in person can require several exchanges online conducted over a period of days. Even instant messaging or chat services can be awkward. The lack of visual cues such as body language, facial expressions, or tone of voice can lead to false starts and miscommunication.

Eileen Abels has proposed a model for communicating effectively online. She examined numerous reference encounters conducted online and identified distinct patterns of communication. Based on these patterns, she suggests that an efficient e-mail reference interview should consist of no more than three stages: an initial problem statement by the patron, a quick response containing a summarization by the librarian, followed by a confirmation by the patron. This need not be the end of the communication, as the answer may not yet be obtained, but if done correctly the interview should be completed at this point. If the interview proves to be more complicated, Abels recommends using another medium (telephone, or in-person) to complete the interaction.[28]

The last two RUSA Behavioral Guidelines stress the importance of involving the patron in the process, and engaging in follow-up. These should be a part of any reference interview. Reference service is more than simply a correct answer. Just as important are what market researchers refer to as "dirty bathroom" factors. Users are unlikely to return to a restaurant with dirty restrooms, even if the food was acceptable. Similarly, research by Joan Durrance has found that 90% of users would not return to an unfriendly librarian, even if they were satisfied with the information obtained.[29] One way to measure the behavioral aspects of your electronic reference service may be to record the number of repeat users. If your patrons are not returning, this could indicate that the service is perceived to be unfriendly.

An assessment focusing on behavioral issues can be conducted a number of ways. An unobtrusive method might involve asking people to pose as users of the service. The responses sent to these users can be examined by library staff to see if proper reference interview and follow-up techniques were used. This technique can be problematic, however, as reference staff may become resentful if they feel they are being spied upon. These issues can be minimized if it is

made clear that the assessments are being done to improve services and not to punish or penalize individual librarians. Finally, the users are the ultimate judges of performance. A survey or interview can be conducted to ask users whether they felt the librarian involved them in the process, or asked follow-up questions.

ONGOING AND FUTURE DEVELOPMENTS
IN ASSESSING ELECTRONIC REFERENCE SERVICES

It is clear from the preceding discussion that the library profession does not lack criteria with which to assess electronic reference services. What is scarce, however, are standards or benchmarks with which to compare one's institution. For example, if your library's policy is to respond to e-mail questions in 24 hours, is this good, bad, or about average? The profession needs more researchers working on developing specific measures that can apply to a variety of library environments.

An excellent model to emulate is provided by the Virtual Reference Desk AskA consortium. The quality standards for the consortium are outlined in the Summer 2000 issue of *Reference and User Services Quarterly*. Members in the consortium agree to provide a service that has the following features: Accessible, Prompt Turnaround, Clear Response Policy, Interactive, Instructive, Authoritative, Private, Reviewed, Referrals, Publicity.[30]

What makes the VRD document outstanding are two features. One, it provides detailed measurement criteria for each element of quality service. For example, "Prompt Turnaround" is defined using specific percentages of questions that should be answered in a defined period of time. Another excellent feature of the standards is that they are flexible. For each element there is a base target required of all members, a current practice that reflects the current level of services offered, and a future goal, which corresponds to the optimal level that members should strive to achieve over time. This built-in flexibility is key. As technology changes and improves, it is important that our service expectations do not remain static.

The future of electronic evaluation offers many exciting possibilities. One interesting endeavor is the Association of Research Libraries New Measures Project. The ARL is investigating the feasibility of establishing ServQual criteria as a de facto national library standard. Briefly, the ServQual criteria are:

- *Reliability*; i.e., ability to perform the promised service dependably and accurately.
- *Assurance*; i.e., knowledge and courtesy of employees and their ability to convey trust and confidence.

- *Empathy*; i.e., the caring, individualized attention the firm provides to its customers.
- *Responsiveness*; i.e., willingness to help customers and provide prompt service.
- *Tangibles*; i.e., appearance of physical facilities, equipment, personnel, and communications materials.[31]

The elements of the ServQual assessment are being modified and tested within ARL Libraries. The proposed new instrument has been named LibQual+. If LibQual+ is widely adopted by ARL libraries, it may lead to the collection of standardized quality measures. These will allow libraries to compare their performance with peer institutions.[32]

Another initiative of interest is the Digital Library Federation's Discussion Group on Usage, Usability, and User Support. The members of this group are surveying libraries to determine what evaluation needs exist. They are seeking to determine the gaps in the research with a goal of devising new research methods to address these information gaps.[33]

Technology offers additional cause to be optimistic about the future. Most interactive software packages include built-in mechanisms for capturing and recording vital transaction data. The software can automatically record information such as the number of interactions, the length of each transaction, and average wait time for a session. The quality of the interaction can be observed as session transcripts are automatically generated. Privacy concerns can be addressed by stripping out names from the record before it is analyzed. Below is an example of statistics and a session transcript from the University of Chicago's online reference service:

**Ask a Librarian Live
Statistics for September 2000.[34]**

Number of Chat Sessions	143*
Total Time in Chat	12:39:42
Average Time in Chat	00:05:18
Average Time in Queue	00:01:00

*This number also includes the practice session that were done to get familiar with the service

**Ask a Librarian Live
Sample Session Transcript.[35]**

Operator Librarian is connected to The Business and Economics Resource Center I-Operator.
Patron is connected to The Business and Economics Resource Center I-Operator.

Librarian: Hello. You have reached the Business & Economics Resource Center at The University of Chicago. How may I help you?

Patron: I was wondering if it is possible to obtain an economics paper titled "Economics of Superstar" written by Sherwin Rosen (UofC Professor) in 1981?
Librarian: Do you know what journal it was published in?
Librarian: Nevermind, it is in American Economic Review.
Patron: You are great, thank you . . . how would I be able to get a copy?
Librarian: You will have to order it via Interlibrary Loan through your company or public library. Licensing restrictions don't allow me to send it to you. The full citation, however, is: Rosen, Sherwin. "The Economics of Superstars" American Economic Review; v71 n5 Dec. 1981, pp. 845-58.
Librarian: Are you a University of Chicago faculty, student or staff member?
Librarian: If you are, I can tell you how to get access to the databases.
Patron: It is actually for my boss, who did get his undergraduate degree from UofC–his name is XXXXXXX.
Librarian: Alumni with current borrowing privileges are eligible but must pay $5 per transaction.
Librarian: The number for Interlibrary Loan is: 773-702-7031
Librarian: Call them. I'm not sure if their service includes possible faxing this article to you; definitely ask.
Librarian: Is there anything else I can help you with?
Patron: I really appreciate all your help, I will give them a call. Thanks again and have a good day.
Librarian: Please click "quit" and fill out the end-user survey (it would help us out!)
Librarian: Thanks!

The transcript and statistics above demonstrate the potential for quantitative and qualitative evaluation of online reference services. What are needed in the future are benchmarks based on the results of research conducted in a wide variety of library settings. More libraries need to evaluate their electronic reference services, and the results must be widely disseminated. Only by sharing our successes (and failures) will the profession move forward. It is hoped that this article has motivated at least a few researchers to begin seriously assessing services at their own institutions.

ENDNOTES

1. Association of Research Libraries (1999), *SPEC Kit #251: Electronic Reference Service,* Washington, DC, p. 7.

2. *Ibid*.

3. Among the formal evaluations conducted are: Lara Bushallow-Wilbur, Gemma DeViney, and Fritz Whitcomb, "Electronic Mail Reference Service: A Study," *RQ* 35 (Spring 1996): 359-371; Ann Bristow, "Academic Reference Service over Electronic Mail: An Update," *College and Research Libraries News* 7 (1995): 460; Julie M. Still and Frank Campbell, "Librarian in a Box: The Use of Electronic Mail for Reference," *Reference Services Review* 21, no.1 (1993): 15-18; Beth A. Garnsey and Ronald R. Powell, "Electronic Mail Reference Services in the Public Library," *Reference & User Services Quarterly* 39, no. 3 (Spring 2000): 245-252.

4. Ferguson, Chris D. and Charles A. Bunge, "The Shape of Services to Come: Values-Based Reference Service for the Largely Digital Library." *College & Research Libraries* 58, no.3 (1997): 257.

5. McClure, Charles R., "User-based Data Collection Techniques and Strategies for Evaluating Networked Information Services," *Library Trends* 42: 592.

6. Miller, Bill, "What's Wrong with Reference: Coping with Success and Failure at the Reference Desk," *American Libraries* 1984: 303-306, 321-322.

7. Evaluation of Reference and Adult Services Committee. American Library Association (1995), *Reference Assessment Manual,* Ann Arbor, Michigan.

8. Whitlatch, Jo Bell (2000), *Evaluating Reference Services: A Practical Guide,* American Library Association: Chicago.

9. For examples of case studies see Whitlatch, Jo Bell, *Evaluating Reference Services: A Practical Guide,* p. 91-97.

10. For a discussion of the various cost-benefit methods, see Chapters 3 and 14 in *The Reference Assessment Manual*.

11. For additional information on focus groups see David L. Morgan & Richard A. Krueger (1998), *The Focus Group Kit,* Sage: Thousand Oaks, Calif.

12. Examples of studies using observational techniques are: Childers, Thomas, "The Test of Reference," *Library Journal* 105: 924-928, and Patricia Dewdney and Catherine Sheldrick Ross, "Flying a Light Aircraft: Reference Service Evaluation from a User's Perspective." *RQ* 34: 217-230.

13. For additional information on conducting surveys see Suskie, Linda (1992), *Questionnaire Survey Research: What Works,* 2nd ed., American Association for Institutional Research: Tallahassee, Fla.

14. Chisman, Janet, Karen Diller, and Sharon Walbridge, "Usability Testing: A Case Study," *College and Research Libraries* 60, no. 6 (1999): 552-569.

15. Garnsey, Beth A., & Ronald R. Powell, "Electronic Mail Reference Services in the Public Library," *Reference & User Services Quarterly* 39, no. 3 (Spring 2000): 245-252.

16. *Reference Assessment Manual,* (1995): p. 7.

17. *Ibid*.

18. Mon, Lorri, "Digital Reference Service," *Government Information Quarterly* 17, no. 3, (2000): 310.

19. Ferguson & Bunge (1997).

20. Richardson, Joanna, "*Ask a Librarian" Electronic Reference Services: The Importance of Corporate Culture, Communication and Service Attitude*. Online, *<http://www.bond.edu.au/library/jpr/ausweb2k/>* Accessed: 12-15-2000.

21. Ferguson and Bunge (1997).

22. Johnston, Pat & Ann Grusin, "Personal Service in an Impersonal World: Throwing Life Preservers to Those Drowning in an Ocean of Information," *The Georgia Librarian* 32 (1995): 45-49.

23. Straw, Joseph E., "A Virtual Understanding: The Reference Interview and Question Negotiation in the Digital Age," *Reference & User Services Quarterly,* 39, no. 4 (Summer 2000): 376-379. Also, Myoung C. Wilson, "Evolution or Entrop: Changing Reference/User Culture and the Future of Reference Librarians," *Reference & User Services Quarterly* 39, no. 4 (Summer 2000): 387-390.

24. Association of Research Libraries (1999), *SPEC Kit #251*.

25. McClure, Charles R. and Cynthia L. Lopata (1996), *Assessing the Academic Networked Environment: Strategies and Options*. Washington, DC, Coalition of Networked Information: 5.

26. Arms, William Y., "Automated Digital Libraries: How Effectively can Computers Be Used for the Skilled Tasks of Professional Librarianship?" *D-Lib Magazine,* 6, no. 7/8 (July/August 2000). Online, *<http://www.dlib.org/dlib/july00/arms/07arms.html>* Accessed: 12-08-2000.

27. Statistic cited in, Gray, Suzanne M., "Virtual Reference Services: Directions and Agendas," *Reference & User Services Quarterly* 39, no. 4 (Summer 2000): 372.

28. Abels, Eileen G., "The E-mail Reference Interview," *RQ* 35, no. 3 (Spring 1996): 345-358.

29. Durrance, Joan C., "Reference Success: Does the 55 Percent Rule Tell the Whole Story?" *Library Journal* 114 (April 15, 1989): 35.

30. Kasowitz, Abby, Blythe Bennett, & R. David Lankes, "Quality Standards for Digital Reference Consortia," *Reference & User Services Quarterly* 39, no. 4 (Summer 2000): 355-363.

31. Cook, Colleen, Fred Heath, Bruce Thompson & Russell Thompson, "The Search for New Measures: The ARL LibQual+ Project–A Preliminary Report," *portal: Libraries and the Academy* 1, no. 1 (2001), 65-74. Online, *<http://muse.jhu.edu/journals/ portal_libraries_and_the_academy/v001/1.1cook.html>*. Accessed: 12-20-2000.

32. For additional information on LibQUAL+ see: *http://www.arl.org/libqual/*.

33. Digital Library Federation, "Usage, Usability, and User Support: Report of a Discussion Group at the DLF Forum on 2 April 2000," Online: *<http://www.clir.org/ diglib/use/useframe.htm>*. Accessed: 12-20-2000.

34. "About Ask-A-Librarian Live," University of Chicago, Business and Economics Resource Center *<http://www.lib.uchicago.edu/e/busecon/aboutask.html>*. Accessed 1-20-2001.

35. "About Ask-A-Librarian Live," University of Chicago, Business and Economics Resource Center *<http://www.lib.uchicago.edu/e/busecon/aboutask.html>*. Accessed 1-20-2001.