Evaluating Reference Services in the Electronic Age

Jo Bell Whitlatch

Abstract

IN AN ELECTRONIC ERA, THE EVALUATION OF REFERENCE and related information services should still be based on the same principles used to evaluate traditional face-to-face reference services and printed reference tools. Traditional research methods—which are surveys and questionnaires, observation, individual and focus group interviews, and case studies—can be utilized very effectively in an electronic environment. However, electronic technologies offer interesting research opportunities not present in the traditional reference environment.

Introduction

At conferences and workshops on evaluating reference services, the most frequent recurring question librarians ask is, "How can the material on evaluating reference services be applied to assessing electronic reference services?" The best answer is, "Take existing methods, determine which will best meet the study goals, and then adapt those methods to the electronic environment."

In any environment, evaluating reference services still requires starting by assessing why reference services are being evaluated and what the organization plans to do with the study results. Before trying to decide how to evaluate electronic services, performance standards that set the level of achievement expected for the service should be explicitly stated. In determining the performance standards to be adopted, the organization must decide what values are crucial. Are members of the organization concerned primarily with

Jo Bell Whitlatch, Reference Librarian & History Selector, San Jose State University, San Jose CA 95192–0028

1. Economics—the cost or productivity of services;
2. The process—aspects of librarian/reference system and user interaction;
3. Resources—books, indexes, databases, staffing levels, equipment, design of physical or electronic environment; or
4. Products/outcomes—information or knowledge that the users obtain.

In an electronic environment the interactions between librarians and users often will no longer be truly face to face. Thus, process standards are the measures that most need to be reviewed in a digital reference environment. Librarian behaviors that are crucial in the reference-desk environment will need to be redefined for remote reference services. Work on redefining process standards has already begun. The Virtual Reference Desk (VRD) project has developed a list of User Transaction Standards to address aspects of librarian/system and user interaction. The standards address several "facets" related to quality: accessible, prompt turnaround, clear response policy, interactive and instructive (Kasowitz, Bennett & Lankes, 2000). Most of these facets address the process standards, rather than standards related to economics, resources, or products/outcomes.

In a remote electronic reference environment, accessibility and prompt turnaround could become dominant in user evaluations. Miwa (2000) used digital reference services features of acknowledgment, responsiveness, and tone of message to represent the process aspects of the reference interaction in a digital environment. She also looks at user situations as part of the process—for example, wording of the request by the user and user's ability to comprehend the message.

Broad goals for the study should be prepared in writing once a reasonable degree of consensus has been achieved on the particular set of standards that an organization wishes to emphasize. After broad goals have been developed, written objectives should be developed for each study goal. The objectives should be measurable so that, at the conclusion of the evaluation, one can identify any gaps between the present level and the desired level of reference service performance.

This present paper discusses how to apply traditional evaluation methods in an electronic reference environment once the study goals and objectives have been determined. Readers desiring additional information on setting performance standards and developing goals and objectives for reference service evaluation may wish to consult *Evaluating Reference Services: A Practical Guide* (Whitlatch, 2000).

All methods have strengths and weaknesses. Depending on the goals and objectives of the study, some methods will be more effective than others. As a general rule, utilizing more than one method is recommended in a single study, because the strengths of one method often compensate for the weaknesses of another. The advantages and disadvantages of the various methods may also change somewhat in an electronic environment. This

paper considers how applying surveys, observation, interviews, and case studies—all traditional evaluation methods used in assessing face-to-face services—presents new opportunities and challenges in assessing electronic reference services.

SURVEYS AND QUESTIONNAIRES

Surveys or questionnaires are methods of directly collecting information on individuals' thoughts, beliefs, attitudes, and opinions, plus objective data, such as education, gender, and income. The survey method has been the most frequent way of assessing traditional reference services. In the past, surveys have been relied upon too heavily because they are the most efficient method of assessing a large group of representative users. Also, for the inexperienced researcher, surveys appear easy to design. The disadvantages, such as obtaining meaningless information from poorly designed questions and the lack of depth of information from standardized responses, are often not appreciated until too late. Another significant problem in using surveys is low response rates, particularly from surveys distributed through the mail. A substantial number of nonrespondents can bias the results; those who choose not to complete the survey might hold very different views from those who do.

Internet questionnaires can be used effectively to survey attitudes and opinions on the quality of reference service related to process (the interaction with the virtual reference service) and products/outcomes (the value of the information obtained). An Internet survey asking for an evaluation of service provided can be sent out within days after the user has received an answer. In contrast to surveys distributed in person at the reference desk or in the library, emailing the questionnaire can also be calculated to allow most users some time to use and further evaluate the information obtained through a reference interaction.

As Zhang (1999) points out, the Web provides new opportunities to conduct survey research more efficiently. Research costs for sending out Internet surveys are relatively low and the turn-around time short compared to conventional mail-in surveys. Also, email can be used effectively to follow up on paper-based surveys (Roselle & Neufeld, 1998). Most responses received in electronic format have been precoded, eliminating transcription errors and saving time and expense. McCullough (1998) notes that Web-based surveys are faster, generate more accurate information, and cost less. He has found that a respondent will typically complete a Web-based survey in about half the time it would take an interviewer to conduct that survey by telephone or in person.

Resolving the technical problems with Internet surveys requires greater technical expertise on the part of the researcher than does research conducted with traditional survey methods. However, services that provide Web survey forms and guidance to assist researchers in designing and de-

veloping Internet surveys are becoming common. Names and Web address-
es for some of the services that have been positively discussed on the Acad-
emy of Management Research Listserv, *rmnet@listserv.unc.edu,* are provid-
ed in the Appendix.

Zhang (1999) also reviews potential problems and concerns related to
Internet-based surveys. One of the greatest strengths of survey research is
the ability to randomly select respondents in a manner that ensures a sam-
ple representative of the target population. In telephone surveys, respon-
dents are randomly selected, but most online poll respondents are self-
selected (Pew Research Center, 1999). The greatest difficulties with Internet
surveys occur when the survey does not reach certain types of respondents
who need to be included in the survey population. Biased samples and re-
turns can be a major problem because certain social groups are underrep-
resented among Internet users.

However, for surveying users of electronic reference services, bias
should be minimal. Respondents must have access to the Internet in order
to use the electronic services; they can presumably access a Web survey form
as well. Some individuals may not have convenient access from their home
or office and may use the service only occasionally from an Internet café
or a library. If these individuals are not identified, this group may be un-
derrepresented. Individuals who do not have convenient access may, as a
whole, be less experienced users of electronic reference services. If these
users are not included in the sample, survey results may not truly represent
the population as a whole. Other means, such as a telephone interview or
mail survey, may be required to obtain responses from them. Finally, if the
purpose of the survey is to collect information from people who do not use
electronic sources, relying upon the Internet as the principal method of
survey delivery will present a very serious problem.

In addition, low response rates are a serious problem with Internet
surveys. In her evaluation of AskERIC, Shostack (2000) observed that us-
ers were either extremely happy or dissatisfied with digital reference ser-
vices. These results suggest that only motivated users are responding. A study
that replicated an earlier study found a disturbing decline in email response
rates: in 1995 the email response rate was 80 percent, but by 1998 it had
fallen to 42 percent (Bachmann, Elfrink & Vazzana, 1999). The research-
ers suggest that the most likely reason for the decline is the respondents'
increased reluctance to respond by email.

Zhang (1999) concludes that the Internet cannot serve as the only
means to collect survey data if researchers need representative returns from
a sample. Schaefer and Dillman (1998) found that giving advance notice
requesting participation generally increases response rates. The Pew Re-
search Center (1999) has tested an interesting approach. Email addresses
were collected from individuals who were called as part of randomly selected
national samples. If these individuals agreed to participate in a future on-

line survey, their email addresses were placed in a pool. Then, in a second phase, a random sample was selected from this pool. Email addresses were used for verification purposes to prevent respondents from taking the survey more than once. McCullough (1998) suggests that the questionnaire be posted on a secure Web site. Respondents can be generated from personal invitations issued by email. He notes that a sufficiently large sample of 300 or 400 respondents can often be completed over a weekend.

In order to apply scientifically tested polling techniques to Internet technologies, Stanford political scientists Douglas Rivers and Norman Nie have created Knowledge Networks. With $42 million in venture capital, they have installed free WebTV devices normally costing $250 each in 40,000 homes selected through random phone calls. Because everyone in the household nineteen or older is involved, there are about 100,000 participants. The homes receive a black box slightly smaller than a VCR, a cordless keyboard, and many instructions. The homes are expected to remain in the survey pool for three years. In exchange for answering brief surveys about once a week, the households receive free Internet access, email, and frequent chances to win prizes. Of those who were asked to join the Knowledge Networks pool, 56 percent agreed—compared with 15 percent of people who usually agree to participate in phone polls. Although the polling is a significant activity, the primary company income comes from consumer research for manufacturers (Konigsmark, 2000).

Zhang (1999) also reports that validity of Internet survey responses can be adversely affected. Unintended participants may respond because of the ease of forwarding email messages to other people. Individuals can respond to a single survey by submitting the same reply many times. Unique case-identification numbers should be assigned to each respondent to control for multiple responses and unintended participants.

Nondeliverable surveys are also a major disadvantage of email. In 1995 and 1998 studies, Bachmann, Elfrink and Vazzana (1999) found that about 20 percent of all emailed surveys were nondeliverable.

Comfort level with the Internet survey form should also be considered. Zhang (1999) found that, while 80 percent of usable replies were received via the Web, 20 percent of respondents chose to complete the survey via postal mail or fax. Internet survey respondents did report problems with the layout of the survey questionnaire on low-resolution monitors, problems going back to previous parts of the questionnaire, problems with printing, and (on computers with low-speed modems) problems with downloading the questionnaire. Users also reported that comments were also more difficult to insert on electronic survey forms than on paper forms. Shostack (2000) also noted a tendency for users to ignore open-ended questions on Internet survey forms. (This problem is not unique to online surveys. In the author's experience, most users completing paper forms also tend to leave open-ended questions blank.) Surveys not conducted by telephone or in-person

interview tend to have rather limited potential to collect qualitative data. An experiment with incentives in the form of cash prizes revealed that, while the overall numbers of respondents did not increase significantly, the number of completed Internet survey questionnaires did rise (Pitkow & Kehoe, 1996).

OBSERVATION

Observational methods collect information on people as they behave in real-life situations. Forms of observation that have been used to assess the quality of reference services include direct observation of the reference interview, observers disguised as patrons asking preassigned questions, self-observation in the form of diaries or journals, recording interviews with audio or videotape, reviewing data collected as part of daily library operations, and examining information on reference transactions collected for another purpose.

Observational methods have been less frequently used than surveys to evaluate reference services, because this method requires a greater investment of staff time. Safeguarding against observational bias also requires training observers thoroughly and may require using more than one observer.

The electronic reference service environment offers some new and exciting opportunities in use of observational methods. Information on electronic reference transactions can be collected and archived as part of ongoing library operations much more easily than can information on traditional reference interviews. Content analysis of these electronic questions should enable us systematically to study the nature of the questions, sources used, and skills required to a much greater extent than is possible in face-to-face reference interactions. The review and analysis of samples from archives of questions and answers provide a practical tool to diagnose problems and improve services.

Studies of email reference questions that use observational techniques are already underway. Garnsey and Powell (2000) examined and classified email reference questions into one of the following categories based on content: (1) ready reference; (2) research question; (3) genealogy; (4) library technology; (5) request for materials; (6) bibliographic verification; and (7) other. Jones, Carter, and Memmott (1999) used a random sample of academic libraries to study the proportion of libraries offering digital reference services and to examine the characteristics of those services. They looked at size of libraries, direct links from library home pages, ways in which users were able to submit questions, FAQ documents, policies, institutional barriers, and the role of type of institutional funding (public vs. private). Shostack (2000) analyzed questions that had been submitted via a question submission form to AskERIC. She found that over 80 percent of users filled out the form completely. Staff were also asked to change the subject line of the response to the topic of the reference query so that questions could be classified by topic.

However, the ease of collecting such information does raise the level of concern about protecting the individual's rights to privacy. The first rule of ethics in research is to do no harm to the participants. In using data for research, particular attention must be paid to protecting the identity of individual users when archiving questions and answers. Access should be restricted to all information that might reveal people's identities. Names and specific information that have the potential to identify individual participants, such as physical descriptions, very detailed demographic information, or identifying events or places, should be removed or modified. Without proper protections, publication of the analysis could harm the morale and self-esteem of reference librarians, staff, and users.

Gray (2000) used observational methods to analyze Web sites of ten large research libraries that provide virtual reference services. The approaches to centralization, placement of the link to reference services on the Web page, use of forms, definition of client base, response times, and question types accepted were analyzed. Observational methods are also useful for testing the effectiveness of different types of answering sources. To compare the effectiveness of print and paper-based reference sources in answering different types of reference questions, Havener (1990) divided 68 reference librarians into two different groups. Members of one group were permitted to use only print tools in their research, while members of the other group could use only online sources to answer the same set of questions. Information recorded varied by question type—for conceptual questions, librarians were asked to record ten relevant citations; for factual questions, librarians were asked to provide only one relevant fact. Time spent was also recorded. In an exploratory study, Janes and McClure (1999) compared the accuracy of answers found in freely available Web sites and traditional print-based sources by asking participating librarians and library school students to answer 12 questions only with resources they were directed to use (either Web or non-Web). Connell and Tipple (1999) gathered ready reference questions that were actually asked by users over a two-week period and then, using AltaVista as a search engine, searched for and examined the accuracy of answers found on the Web.

Observational methods are useful in determining the difficulty that users encounter with online reference tools. Chisman, Diller and Walbridge (1999) advertised for volunteers who were paid ten dollars for their participation. A usability test was designed to determine how easily users could navigate a Web catalog and whether they understood what they were seeing. Observers recorded the search strategy, comments made by the participants, observations about the participants' responses, success, and the time needed to complete the task.

Unobtrusive observation methods can also be used effectively in an electronic world. Reference questions can be prepared and answers determined for factual types of questions. Graduate students or others who are posing

as users with questions can query both commercial and non-profit "ask a question" services. Results can be analyzed by such factors as response time, accuracy or quality of answer, tone of message, ease of submitting the question, and observations on whether people would return to the site again.

INDIVIDUAL INTERVIEWS AND FOCUS GROUP INTERVIEWS

Interviews are an appropriate method for collecting information on how people interpret their world, describe their experiences, and articulate their attitudes, perspectives, concerns, and values. Despite the potential for gathering in-depth information, interviews have been less frequently used than surveys because of the expense and time required. As is the case with observational methods, interviewers must be thoroughly trained to avoid bias. The management and coordination of scheduling for either individual or group interviews can be extremely time consuming. Coding and analyzing the data also require considerable time.

Interviews of both users and librarians are also possible in the digital reference service environment. Interviewers can use Web-based survey forms to record the results of interviews efficiently. However, users will probably be harder to reach than in-person users of reference-desk services. Marketers have begun to use online focus groups; chat technology with these methods could certainly be adopted for users of electronic reference services. While online focus groups do not allow moderators to observe how people are interacting, benefits include no geographic barriers, lower costs, more rapid turn-around time, and the possibility that participants may be more open because of the greater anonymity provided by chat rooms (Maddox, 1998).

Conventional focus groups can also be used effectively to evaluate digital reference services. By reaching out to user groups in the community (teenagers at risk, small business organizations, etc.) or distance learning communities in an academic setting, participants can be recruited to assess their experience with digital reference services. Food or some other small gift of appreciation and a convenient location will encourage participation.

CASE STUDIES

Case studies use a combination of assessment methods to analyze services in one or in a limited number of situations. Case studies have been used to assess new reference services or products. Combining the different methods will enrich study findings significantly, but will also increase the time required to conduct the study and analyze the information collected. Results generally cannot be applied to other situations.

Case studies have great potential to improve our understanding of the quality of digital reference services. Using information provision in a hospital setting, Barcellos (2000) is studying user intermediary interactions through use of organizational publications, site observations, transaction

logs, and interviews of both users and intermediaries. A case study of the Internet Public Library Reference Division examined unanswered questions to determine why they were not being answered and to generalize about the difficulties associated with providing reference services via the Internet (Ryan, 1996). White (1999) has developed a framework for evaluating electronic question-answering services that involves World Wide Web inspection, perusal of publicly available policy documents, and personal contact via email and/or interviews with service administrators.

CONCLUSION

Several years ago, James Rettig (1996) observed that many of the criteria used for evaluating printed reference resources have analogs in the digital world: for example, authority, accuracy, level or audience, and content. Standards and methods used for evaluating traditional reference services also have many analogs in the world of digital reference. Standards and criteria related to economic considerations, the reference process, reference resources, or products or service outcomes will still be important in an electronic world. Traditional methods of survey, observation, interview, and case study remain useful.

Case studies that focus on evaluating experimental digital reference services and employ a variety of research methods may have the greatest promise to enhance our knowledge. Case studies have the potential to improve our knowledge of both the effectiveness of digital reference services and the combination of methods best suited to evaluate them. Over time, the profession should, through the effective use of case studies, be able to build a guide to best practices, not only for digital reference services, but also for the methods necessary to assess and continually improve these services.

Results of initial studies of digital reference services and the now well-known phenomenon of declining business at many reference desks also suggest that these studies should be used to analyze future directions in reference practice. Studies (Connell & Tipple, 1999; Janes & McClure, 1999) indicate that freely available Web materials may serve as well as traditional ready reference tools for answering many of the common types of queries received at reference desks. For most users, convenience is first. The expert in-person assistance a librarian might provide is becoming comparatively less convenient than it once was, when the alternative source is the Web. Many users will love the convenience and be satisfied with "good enough." Others will find it more convenient to take advantage of remote ready reference services, which will probably be supported by a relatively small amount of funding or reference librarians from each local library.

As the demand continues to shift away from the reference desk, libraries have the opportunity to establish much more active outreach programs. The public and administrators may come to view reference librarians as less

essential than in past times. While libraries still have reference librarians, shifting patterns of user demands for reference services provide libraries with opportunities to emphasize different strategies to connect library materials with users. Libraries may develop a stronger role in the community in promoting information competencies through partnerships with community service agencies or, within the academic community, with faculty engaged in critical thinking and writing courses.

Changes in strategy would also have implications for professional education. Marketing skills that are essential for developing active outreach programs, as well as instructional skills, may need to become a major part of the core curriculum in every library school. One of the essential marketing skills is evaluation and improvement of outreach efforts. Perhaps the day will come when all librarians engaged in professional practice will receive, as part of their professional education, in-depth understanding and experience in developing and applying survey, observation, interview, and case-study methods so that reference librarians might change, survive and prosper in the new electronic age.

REFERENCES

Bachmann, D. P.; Elfrink, J.; & Vazzana, G. (1999). E-mail and snail mail face off in rematch. *Marketing Research, 11*(4), 10–15.

Barcellos, S. (2000). Understanding intermediation in a digital environment: An exploratory case study Paper presented at Facets of Digital Reference: The Virtual Reference Desk 2nd Annual Digital Reference Conference. Summary available from http://www.vrd.org/.

Chisman, J.; Diller, K.; & Walbridge, S. (1999). Usability testing: A case study. *College & Research Libraries, 60*(6), 552–569.

Connell, T. H., & Tipple, J. E. (1999). Testing the accuracy of information on the World Wide Web using the AltaVista search engine. *Reference & User Services Quarterly, 38*(4), 360–368.

Garnsey, B. A., & Powell, R. R. (2000). Electronic mail reference services in the public library. *Reference & User Services Quarterly, 39*(3), 245–254.

Gray, S. M. (2000). Virtual reference services: Directions and agendas. *Reference & User Services Quarterly, 39*(4), 365–373.

Havener, W. M. (1990). Answering ready reference questions: Print versus online. *Online, 14*(1), 22–28.

Janes, J.; Carter, D.; & Memmott, P. (1999). Digital reference services in academic libraries. *Reference & User Services Quarterly, 39*(2), 145–150.

Janes, J., & McClure, C. R. (1999). The Web as a reference tool: Comparisons with traditional sources. *Public Libraries, 38*(1), 30–39.

Kasowitz, A.; Bennett, B.; & Lankes, R. D. (2000). Quality standards for digital reference consortia. *Reference & User Services Quarterly, 39*(4), 355–363.

Konigsmark, A. R. (2000), High tech joins polls. *San Jose Mercury News*, November 12 B1, B5.

Maddox, K. (1998). Virtual panels add real insight for marketers: Online focus-group use expanding. *Advertising Age*. Retrieved August 10, 2000 from http://adage.com/interactive/daily/archives/id199806.html.

McCullough, D. (1998). Web-based market research ushers in a new age. *Marketing News, 32*, 27–28.

Miwa, M. (2000). User situations in digital reference service: An evaluation of the AskERIC Q & A Service. Paper presented at Facets of Digital Reference: The Virtual Reference Desk 2nd Annual Digital Reference Conference. Summary of presentation from http://www.vrd.org/.

Pew Research Center. (1999). A survey methods comparison: Online polling offers mixed results. Retrieved March 19, 2000 from http://www.people-press.org/onlinerpt.htm.

Pitkow, J., & Kehoe, C. (1996). GVU's 6th WWW User Survey. Paper posted on the World Wide Web. Retrieved January 25, 1997 from http://www.cc.gatech.edu/gvu/user_surveys/survey-10-1996/#exec.

Rettig, J. (1996). Beyond "cool": Analog models for reviewing digital resources. *Online.* Retrieved December 30, 1998 from http://www.onlineinc.com/onlinemag/SeptOL/rettig9.html.

Roselle, A., & Neufeld, S. (1998). The utility of electronic mail follow-ups for library research. *Library and Information Science Research, 20*(2), 153–161.

Ryan, S. (1996). Reference service for the Internet community: A case study of the Internet Public Library Reference Division. *Library & Information Science Research, 18*(3), 241–259.

Schaefer, D. R., & Dillman, D. A. (1998). Development of a standard e-mail methodology. *Public Opinion Quarterly, 62*(3), 378–379.

Shostack, P. L. (2000). Identifying users' needs. Paper presented at Facets of Digital Reference: The Virtual Reference Desk 2nd Annual Digital Reference Conference. Summary of presentation available from http://www.vrd.org.

Virtual Reference Desk. http://www.vrd.org/training/facets.html.

White, M. D. (Ed.) (1999). *Analyzing electronic question/answer services: Framework and evaluations of selected services* (Report No. CLIS-TR-99–02). College Park, MD: College of Library and Information Services, University of Maryland. ERIC Document Reproduction Service No. ED 433019.

Whitlatch, J. B. (2000). *Evaluating reference services: A practical guide.* Chicago: American Library Association.

Zhang, Y. (1999). Using the Internet for survey research: A case study. *Journal of the American Society for Information Science, 51*(1), 57–68.

APPENDIX: SURVEY ASSISTANCE ON THE WEB

Internet Survey Solutions

http://www.clearpicture.com/Survey_Solutions.htm

Web-based Clear Picture survey system.

Research Internet Advertising Resource Guide

http://www.admedia.org/internet/research.html

Annotated entries for research firms, online surveys, virtual focus groups, survey software.

Survey Select

http://www.surveyselect.com/

Samples of the Saja software product available for viewing on the Web site.

Zoomerang Create Surveys

http://www.zoomerang.com/build_preview/new-survey.zgi?1182

Survey templates for business, community, personal/social, and education.