Library & Information Science Research 30 (2008) 291–297

Contents lists available at ScienceDirect

Library & Information Science Research

Bibliometric analysis to identify core reference sources of virtual reference transactions

Pnina Shachaf *, Debora Shaw

School of Library and Information Science, Indiana University, 1320 E 10th Street, Bloomington, IN 47405-3907, USA

ARTICLE INFO

Article history:
Available online 2 October 2008

ABSTRACT

As the use of electronic reference sources becomes commonplace, virtual reference services are expanding in scope, modes, and popularity. Simultaneously, reference practices are evolving as well. One concept that may be challenged by these trends is the notion of the core reference collection. What are the sources that form this core collection, and what are its characteristics? Are similar sources used to answer users' questions in virtual and traditional reference? How do core collections of public and academic libraries differ? An analysis of 1851 e-mail and chat reference transactions from public and academic libraries reveals that the notion of a core reference collection persists in the world of virtual reference services. In both types of libraries, responses to patrons showed a skewed bibliographic distribution; librarians used a small group of sources to answer most of the questions. Almost all sources used were electronic. Academic libraries tended to make greater use of fee-based sources, but public libraries more often used sources freely-available on the Web.

1. Introduction

"An exhaustive list of libraries offering … virtual reference services is nowadays akin to a list of libraries that offer telephone service" proclaimed LISWiki (2007). New models, such as "Librarian with a Latte" at the University of Michigan (Carlson, 2007), are emerging. Reference librarians are often expected to move seamlessly between in-person, e-mail, and chat interactions. These librarians expand reference practices by using a wide range of modes of communication and diverse information sources. Researchers have focused attention on the quality of these virtual reference services, producing a flood of articles and books. They have also identified guidelines that provide standards and best practices to enhance the quality of virtual reference services. These studies have addressed the reference encounter, librarian behaviors, patron demographics, and the type and volume of questions. However, little research to date has focused on the sources used. It is possible that virtual reference will alter the nature or even the existence of the core reference collection. On the other hand, it is possible that the sources used to answer users' questions in virtual and traditional reference are similar and that librarians rely on the same core collection.

Understanding the resources used in virtual reference services can help to prepare, educate, and train the LIS students, practicing librarians, and library staff who will provide these services. Such knowledge may also be helpful in educating library users and developing expert systems to support virtual reference services. With the high costs for some electronic sources and the ready availability of some Web-based alternatives, knowing which sources reference librarians actually use also has implications for collection development in support of reference services.

Several questions arise in considering these problems:

- Are there similarities between the sources used to answer users' questions in virtual and traditional reference?
- Is the concept of a core/ready-reference collection useful in online services?
- How extensive is the use of electronic sources? of fee-based sources?
- Are there differences in the sources used in answers to virtual reference questions in academic and public libraries?

2. Literature review

2.1. Reference collections

A reference collection is inherently practical: it exists to provide easy access to frequently used sources of information. Keenan and Johnston (2000) described ready-reference material as "reference works that can provide information quickly and easily" (p. 208). Sometimes these collections are developed carefully: Gotsick, Friedman, and Smith (1979) reported a close collaboration between the primary user of a small medical library and an expert consulting librarian to develop their core reference collection. Berkov and Morganstern (1990) used recommendations from experienced library staff as the primary source for a set of core reference sources. In other cases, core reference collections have been accepted as given (Nichols, 1993).

* Corresponding author.
E-mail addresses: shachaf@indiana.edu (P. Shachaf), shawd@indiana.edu (D. Shaw).

0740-8188/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.lisr.2008.04.002

From a bibliometric perspective, one assumes that information on a topic will be widely scattered among many sources, but a small number of sources will provide most of the information (Bradford, 1950). Zipf's (1949) "principle of least effort" would encourage the sensible librarian to place frequently used sources close at hand. Librarians' practices reflect these understandings, although there is evidence that reference collections, like others, benefit from frequent review. Bradford (2005) examined the use of print reference collection in a small college library. She found that less than 10% of the collection was used even once in an academic year. Colson (2007) found that 12% of a small academic library's reference collection was heavily used, 17% was moderately used, and 36% was lightly used.

Most recent studies of the core reference collection reflect the transition from print to digital publication. Reference librarians have adopted and improved access to Web-based sources (Smith, 2001). Bradford, Costello, and Lenholt (2005) analyzed the sources college reference librarians used in answering questions at the reference desk. Databases and other librarians were the most frequent sources (each accounted for 24% of the answers); the library catalog accounted for 15%, and Web pages developed by the library accounted for 12%. Most requests (75%) were answered by a single source. The ARL libraries responding to Tenopir and Ennis's (2002) survey estimated that the local online catalog (source for 29% of answers) and commercial online databases (26%) provided the answers to most of their reference questions. Print (18%) ranked third, and the Web (16%) ranked fourth (p. 272).

2.2. Virtual reference service

Research and evaluation of virtual reference services have focused on the context/marketing of the service, how information seekers initiate requests for information, how librarians respond, interactions between information seeker and librarian, satisfaction of the information seeker, archiving records of the transaction, and other practical issues in providing service (see, for example, McClure, Lankes, Gross, & Choltco-Devlin, 2002; Pomerantz, 2005; Shachaf, 2007; Virtual Reference Desk Network, 2003a, b; White, 2001). Numerous studies have also considered user demographics and the types of questions asked (e.g., Bolander, Connaway, & Radford, 2006; Pomerantz, 2005).

Analyzing the sources consulted in responding to requests for information may be included in digital reference evaluation, but it is usually a relatively minor factor in assessing the quality of service. The Reference and User Service Association (RUSA) of the American Library Association (ALA) specified in one of the four main areas of the guidelines that:

> As an effective searcher, the librarian: … constructs a competent and complete search strategy. This involves: … 4.2 Identifying sources appropriate to the patron's need that have the highest probability of containing information relevant to the patron's query … [The librarian] 4.9 Offers pointers, detailed search paths (including complete URLs), and names of resources used to find the answer, so that patrons can learn to answer similar questions on their own. (American Library Association, Reference and User Services Association, 2004, item 4)

Similarly, the Virtual Reference Desk Network (2003a, b) recommended that digital reference services "promote information literacy by responding with detailed search paths and sets of resources that either provide answers or allow users to investigate on their own" (item 5). Pomerantz (2005) mentioned "searching resources" as one of the essential processes in providing chat reference service (p. 1298). White (2001) suggested using standard resources as a means of maintaining quality (p. 229). McClure et al. (2002) considered "sources used per question ... an important descriptive measure that will have a decided impact on broad-based decisions concerning allocation of resources both for the digital reference service and for the library as a whole" (p. 29). In her chapter on maintaining and building reference skills and knowledge, Kovacs (2007) included 15 assignments, including developing functional knowledge of ready-reference sources of information and awareness of the best or core reference sources available in print and electronic formats.

One of the few works to consider a virtual ready-reference collection (VRRC, in their terminology) was prepared by Mizzy and Mahoney (2002). Focusing on chat reference, they promoted the development and constant review of "a *shared personal space* of the chat librarians ... a VRRC [that] can save keystrokes and enhance reference performance" (p. 70). Their steps for collection development included reviewing logs of e-mail and chat reference interactions to identify sources that had been used effectively. They also emphasized the importance of local information sources, suggesting that the most important would be the local OPAC. Their article included 15 "exemplary virtual ready-reference guide sites" and 27 selection aids, ranging from about.com to WWW Virtual Library. Analyzing e-mail and chat reference transactions is a useful approach to identifying sources used in virtual reference services; it is time to move beyond single-institution studies and try to identify typical practices.

3. Procedures

Researchers analyzed the content of 1851 e-mail and chat transactions from public and academic libraries to identify the sources librarians used in responding to user queries.

3.1. Sources

This study reviewed records of virtual reference transactions from two sources: The "Ask A Librarian" service of the Indiana University Libraries in Bloomington (IU) and the QuestionPoint (QP) chat reference service (http://www.questionpoint.org/) offered by the Library of Congress and OCLC, Inc. The Indiana University Libraries Ask A Librarian service "is intended for the students, faculty and staff of Indiana University, Bloomington"; it also handles questions from anyone about the university and unique resources of the university libraries (Indiana University Libraries, 2008). QP offers online e-mail and chat reference services through a global network of cooperating libraries. Just a year after its creation by the Library of Congress and OCLC, Inc., QP service was offered through 1000 libraries–primarily public and academic–in 20 countries (O'Leary, 2003). In 2007 it handled its 3 millionth question (OCLC, Inc., 2007).

The data include 1851 e-mail and chat transactions that were answered between December 2005 and October 2006. These include 1351 IU e-mail transactions from July 2006 through October 2006 (July — 230; August — 300; September — 383; October — 398; approximately 11 transactions per day) and 450 randomly selected QP chat transactions (50 transactions per month from December 2005 through August 2006). All personal identifiers were scrubbed from the transactions before data analysis began. Researchers numbered each transaction and grouped them by month.

3.2. Methods

Using content analysis, researchers identified the reference sources that librarians mentioned in their responses to the requests. This enabled researchers to identify the sources used, their types, and the frequency with which they were used.

The coding scheme was developed from the data in order to identify information to help to address the research questions (Allen & Reser, 1990; Miles & Huberman, 1994). In addition, once a preliminary set of codes had been developed, the researchers consulted previous

Table 1
Number of transactions, answers that mentioned sources, titles, and unique titles

Period	Data set	Number of transactions	Number (and percent) of answers that mentioned sources	Number of titles	Number of unique titles
July 2006	IU	230	151 (65%)	185	110
August 2006	IU	300	121 (40%)	164	85
September 2006	IU	383	154 (40%)	197	93
October 2006	IU	398	193 (48%)	304	146
TOTAL IU		1311	619 (47% of answers)	850	327 (38% of titles)
December 2005	QP	50	32 (64%)	49	48
January 2006	QP	50	32 (64%)	92	87
February 2006	QP	50	38 (76%)	66	63
March 2006	QP	50	29 (58%)	58	55
April 2006	QP	50	34 (68%)	63	59
May 2006	QP	50	33 (66%)	79	73
June 2006	QP	50	36 (72%)	63	59
July 2006	QP	50	41 (82%)	91	90
August 2006	QP	50	34 (68%)	62	58
TOTAL QP		450	309 (70% of answers)	623	545 (87% of titles)
Total all		1761	935 (53%)	1473	872 (59%)

IU = "Ask A Librarian" service of the Indiana University Libraries in Bloomington.
QP = OCLC QuestionPoint.

reference collection studies (Bradford, 2005; Bradford et al., 2005), and a virtual reference question typology (Shachaf, Meho, & Hara, 2007), examining the applicability of the codes and schemes developed in these studies to the data. This made it possible to adjust the coding scheme, assuring that the codes were exhaustive, mutually exclusive, and related to the research questions (Allen & Reser, 1990). The codes were organized under a conceptual structure (Miles & Huberman, 1994) that provided several broad categories (e.g., Type of source; Question type) and sub-categories (e.g., Reference sources as a sub-category for Type of source). The scheme was tested on a sample of transactions from the data and then finalized. The researchers followed Miles and Huberman's (1994) suggestion to "check coding about two-thirds of the way through the study" (p. 64). Once a significant portion of the data was coded, the researchers reexamined the scheme and eliminated several codes that had not been assigned to any transaction (e.g., the question type—a request for reproduction of materials). Other codes (e.g., Government website) were added, and all the transactions were checked and coded for these additions. Appendix 1 presents the final, 34-item coding scheme.

One person coded all the transactions; a second coder coded some of the transactions to check reliability. Two iterations were required to produce the desired inter-coder reliability above 90% (Miles & Huberman, 1994). After review, the inter-coder agreement was 98% and Cohen's Kappa was .85.

Table 2
Top 20 sources in IU transactions (July 2006–October 2006)

Title	Frequency (% of top 20)
Indiana University	484 (76%)
Academic Search (EBSCO)	33 (5%)
WorldCat	22 (3%)
Dissertation Abstracts	14 (2%)
Lexis-Nexis Academic	13 (2%)
Web of Science	7 (1%)
Google™ Scholar	6 (1%)
Indiana Daily Student Online	6 (1%)
New York Times	6 (1%)
Wikipedia	6 (1%)
ABI/INFORM Global Suite	5 (1%)
Chicago Manual of Style Online	5 (1%)
ERIC	5 (1%)
Indiana State Library Ask a Librarian	5 (1%)
Literature Resource Center	4 (0.5%)
Monroe County Public Library	4 (0.5%)
PsychINFO	4 (0.5%)
AbeBooks	3 (0.5%)
Biography Resource Center	3 (0.5%)
Factiva	3 (0.5%)
Total	638

Table 3
Top 20[a] sources in QP transactions (December 2005–August 2006 sample)

Title	Frequency (% of top 20)
Library Web pages	83 (36%)
Local catalogs	57 (25%)
Google™	19 (8%)
Wikipedia	19 (8%)
Answers.com™	6 (2%)
EBSCO	6 (2%)
Amazon	5 (2%)
About.com®	5 (2%)
Yahoo!®	3 (1%)
WorldCat®	3 (1%)
PubMed	2 (1%)
People's Network	2 (1%)
Medline Plus®	2 (1%)
Library of Congress Catalog	2 (1%)
JSTOR	2 (1%)
InfoTrac®	2 (1%)
How Stuff Works	2 (1%)
CIA World Factbook	2 (1%)
California Law. Find California Code	2 (1%)
Australian Bureau of Statistics	2 (1%)
Total	226

[a] Because 14 sources have the same frequency (2), the table includes 23 titles.

4. Findings

Table 1 provides the number of transactions per month from each data set. Fifty-three percent of the transactions (935 out of 1761) involved answers that mentioned sources. In these transactions 1473 titles were mentioned, of which 872 (59%) were unique. A higher percentage of the Question Point (QP) than Indiana University (IU) answers mentioned sources (70% to 47%). The percentage of unique titles mentioned was also higher in the QP transactions (87%, compared to 38% for IU). Two-way contingency tables analyses were conducted to evaluate whether the number of answers that mentioned sources and the number of unique sources among the sources mentioned differed between the two data sets (IU and QP). Significantly more QP transactions than IU transactions mentioned sources. The number of answers that mentioned sources and the data sets were significantly related Pearson $\chi 2$ (1, $N=3072$) = 10.32, $p<.01$, Cramer's $V=.058$. The number of unique titles and the data sets were significantly related Pearson $\chi 2$ (1, $N=2323$) = 92.72, $p<.01$, Cramer's $V=.2$. More unique titles were mentioned in the QP than the IU transactions.

The 20 sources mentioned most frequently in the IU transactions are listed in Table 2 and those for QP transactions are listed in Table 3.[1]

[1] IU Web pages with unique URLs were counted as independent titles to generate data for Table 1, but all the IU pages were grouped under one category in Table 2. Among the 484 IU Web pages, 62 were created by the university and 422 by the library; this included 175 uses of the library catalog. In the QP transactions, various library pages and catalogs were counted as independent titles in Table 1; in Table 2 all the library Web pages and catalogs were grouped together. Among the 140 library Web pages and catalogs, 57 were local catalogs.

Both tables indicate a skewed bibliometric distribution, with a few sources that were heavily cited and a long tail of other sources. The 20 most frequently mentioned sources in the IU transactions accounted for 75% of the titles in the answers that mentioned sources (638 out of 850). The 20 most frequently mentioned sources in the QP transactions accounted for 36% of the titles in the answers that mentioned sources (226 out of 623). A two-way contingency table analysis revealed that the IU responses relied significantly more often on the top 20 sources: Pearson $\chi 2$ (1, N=1473)=92.72, p<.01, Cramer's V=.389.

Among the most frequently mentioned sources on both lists of top 20 sources were (see Tables 2 and 3): Wikipedia, Google™, EBSCO, WorldCat®, and library Web pages and catalogs. Unique sources among the 20 most frequently used sources in the IU transactions (that were used at least six times) were various library and university Web pages, a few commercial databases (such as Dissertation Abstracts, Lexis-Nexis® Academic, and Web of Science®), and the *Indiana Daily Student Online* (newspaper). Unique sources among the 20 most frequently mentioned sources in the QP transactions (that were cited more than twice) included Answers.com™, Amazon.com, Yahoo!®, and About.com®.

A follow-up examination of the top 20 sources in the QP transaction by type of library revealed that library catalogs and library Web pages were by far the most frequently used sources in answering questions in both types of libraries (Table 4). These were followed by Wikipedia in the public libraries' transactions and Google in both types of libraries. The relatively frequent use of WorldCat, JSTOR, EBSCO, and APA style in the academic libraries' transactions was similar in the IU transactions. Other sources mentioned in the QP public libraries' transactions but not in the QP academic libraries or IU transactions included freely-available Web sources such as About.com, Amazon.com, Answers.com, and Yahoo!.

Researchers searched WorldCat for all resources mentioned in the transactions and recorded their Library of Congress (LC) classification numbers. The first three sections of Table 5 show the frequencies of use by classification. The majority of the resources (i.e., library catalogs, library Web pages, and other freely-available websites) were not cataloged in WorldCat and do not have LC call numbers; therefore, they were not included in this table. The frequency of sources used in any call number ranged from 1 to 36 in the full set of transactions; in the IU transactions it ranged from 0 to 30 (0–21%) and in the QP transactions from 0 to 17 (0–16%). The A (general works) classes had the highest overall percentage (14.5% of all resources). In the IU transactions, the highest percents were for the Z (bibliography/library science) classification (26%), the A classification (19%), and the H (social sciences) classification (14%). In the QP transactions, the greatest use was in the A (16%), H (14%), and K (law) (9%) classifications.

Table 6 summarizes the data on reference sources and types of questions in the IU transactions. Table 7 provides similar information for the QP transactions. The number of sources used to answer a question averaged 1.4 in the IU transactions and 1.6 in the QP transactions. The total number of sources per query ranged from one to nine in both sets of data. Most of the sources in both sets of data

Table 4
Top 20[a] sources in QP transactions: academic vs. public libraries

Title	Frequency (% of top 20)	
	Public library	Academic library
Library web pages	48 (34%)	30 (48%)
Local catalogs	26 (18%)	22 (35%)
Wikipedia	17 (12%)	–
Google™	16 (11%)	2 (3%)
About.com®	5 (3%)	–
Amazon.com	5 (3%)	–
Answers.com™	5 (3%)	–
Library of Congress Catalog	3 (1%)	–
Yahoo!®	3 (1%)	–
Australian Bureau of Statistics	2 (1%)	–
EBSCO	2 (1%)	2 (3%)
How Stuff Works	2 (1%)	–
Medline Plus®	2 (1%)	–
People's Network	2 (1%)	–
WorldCat®	–	2 (3%)
JSTOR	–	2 (3%)
APA style	–	2 (3%)
Total	138	62

[a] Only sources that were used at least twice in the transactions are included in this list: 14 in the public library transactions and 7 in the academic library transactions.

Table 5
Frequencies of LC classification numbers

Call no.	IU		QP		Total		Bradford (2005)	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
A	19	13.4	17	15.9	36	14.5	1329	5.2
B	8	5.6	6	5.6	14	5.6	1173	4.5
C	5	3.5	4	3.7	9	3.6	413	1.6
D	6	4.2	5	4.7	11	4.4	1128	4.4
E	8	5.6	2	1.9	10	4.0	1004	3.9
F	2	1.4	3	2.8	5	2.0	411	1.6
G	4	2.8	6	5.6	10	4.0	740	2.9
H	14	9.9	15	14.0	29	11.6	5494	21.3
J	1	0.7	5	4.7	6	2.4	1299	5.0
K	1	0.7	10	9.3	11	4.4	2082	8.1
L	12	8.5	3	2.8	15	6.0	703	2.7
M	3	2.1	0	0.0	3	1.2	423	1.6
N	0	0.0	1	0.9	1	0.4	427	1.7
P	13	9.2	7	6.5	20	8.0	4716	18.3
Q	3	2.1	6	5.6	9	3.6	2411	9.3
R	3	2.1	6	5.6	9	3.6	372	1.4
S	0	0.0	3	2.8	3	1.2	121	0.5
T	10	7.0	6	5.6	16	6.4	255	1.0
U	0	0.0	1	0.9	1	0.4	162	0.6
Z	30	21.1	1	0.9	31	12.4	1125	4.4
Total	142	100	107	100.0	249	100	25,788	100

Table 6
Frequencies of codes: IU transactions

Code			Total (percent of 850 sources)
Fee			202 (23%)
Number of sources per message	Range		1–9
	Average		1.35
Electronic source			823 (96%)
URL provided			484 (56%)
Reference sources	Catalog	Local catalog	174 (20%)
		WorldCat®	22 (2%)
		Other	4 (.5%)
	Other reference sources	Databases and indexes	167 (19%)
		Census	4 (0.5%)
		Encyclopedia	3 (0.5%)
		Library directory/Web pages/services/policies	262 (30%)
		University URL	49 (5%)
	Non-reference sources	Book	30 (3%)
		Journal	14 (1%)
		Newspaper	7 (0.5%)
		Government website	14 (1%)
		International government website	2 (0.5%)
		Outside URL	115 (13%)
		Other	8 (0.5%)
Type of question		Known item	259 (30%)
		Topical question	234 (27%)
		Technical problem	86 (10%)
		Search instructions	47 (5%)
		Citing instructions	21 (2%)
		Verification	22 (2%)
		Non-reference	60 (7%)
		Other	150 (17%)

Table 7
Frequencies of codes: QP transactions

Code			Total (percent out of 623)
Fee			42 (6%)
Number of sources per message		Range	1–9
		Average	1.6
Electronic source			600 (96%)
URL provided			524 (84%)
Reference sources	Catalog	Local catalog	47 (7%)
		WorldCat®	3 (0.5%)
		Other	10 (1%)
	Other reference sources	Databases and indexes	40 (6%)
		Census	2 (0.5%)
		Encyclopedia	7 (1%)
		Library directory/Web pages/ services/policies	77 (12%)
		University URL	4 (0.5%)
	Non-reference sources	Book	15 (2%)
		Journal	5 (0.5%)
		Newspaper	6 (0.5%)
		Government website	18 (2.8)
		International government website	6 (0.5%)
		Outside URL	376 (60%)
		Other	3 (0.5%)
Type of question		Known item	56 (9%)
		Topical question	341 (54%)
		Technical problem	19 (3%)
		Search instructions	38 (6%)
		Citing instructions	9 (1%)
		Verification	3 (0.5%)
		Non-reference	9 (1%)
		Other	146 (23%)
Type of library		Academic (out of 450)	112 (24%)
		Public (out of 450)	322 (71%)
		Unclear (out of 450)	16 (3%)

were electronic (IU 93%, QP 96%). Librarians provided URLs for these sources most of the time in the QP transactions (84%) and about half of the time in the IU transactions (55%). A two-way contingency table analysis demonstrated that the QP transactions mentioned multiple sources significantly more often than the IU transactions: Pearson $\chi 2$ (1, $N=1473$) = 122.82, $p<.01$, Cramer's $V=.289$. Most of the sources were freely-available on the Web; IU transactions used a higher percentage of fee-based sources (24%) than did the QP transactions (7%). This difference was statistically significant: Pearson $\chi 2$ (1, $N=1473$) = 75.37, $p<.01$, Cramer's $V=.226$. Further analysis of the QP transactions revealed that fee-based sources were rarely used in any transactions: public libraries (5%, 33 out of 574), academic libraries (4%, 8 out of 181), and the other transactions (3%, 1 out of 26).

The IU transactions primarily involved reference sources (78%). Among the reference sources, the most heavily used were library Web pages (38% of the reference sources) and catalogs (29% of the reference sources). QP transactions relied primarily on non-reference sources (69%), and outside URLs were the most frequently used (87% of the non-reference sources). Similar to the IU transactions, most of the QP transactions using reference sources relied on library Web pages (40% of the reference sources) and catalogs (31% of the reference sources).

The most frequent requests in the IU transactions were for known items (29%), followed by topical questions (26%). The QP transactions asked topical questions (57%) most frequently, followed by requests for known items (16%). Fig. 1 shows the distribution of question types.

5. Discussion

Core collections can be identified by analyzing virtual reference transactions. There is overlap in the top 20 sources used in both IU and QP transactions (Tables 2 and 3). Variations between the lists likely reflect in part the different groups of users (academic in the IU transactions vs. mostly public in the QP transactions). The variations between public and academic libraries in the QP transactions are similar to the differences and similarities between the IU and QP transactions (Table 4). However, although the sources used by academic and public libraries varied, library catalogs and library Web pages were the most frequently used in both settings. Google and EBSCO also appeared on both lists. The sources that were unique to the academic list were primarily fee-based indexes and databases; the unique sources on the public list were freely-available Websites.

The vast majority of the sources used in the IU and QP transactions were electronic (96% in each of the two data sets). Bradford et al. (2005) reported that online sources were used to answer nearly 60% of their questions. It is possible that the differences result from the time difference between the two studies: Bradford et al. collected data in 2002, and the data reported here originated primarily in 2006. It is possible that electronic sources are used frequently because their scope, range, quality, utility, and availability have reached a critical point. Also, it is possible that the medium of interaction affects the medium of sources used; Bradford et al. analyzed transactions at the reference desk rather than virtual interactions. In any case, increasing use of electronic sources over the years in both academic and public libraries was expected and has been documented (e.g., Havener, 1990; Tenopir & Ennis, 2002).

Lists of top 20 sources revealed that in the IU transactions, fee-based indexes, and databases were used more often than in the QP transactions (23% vs. 6%, respectively); the QP transactions made much greater use of freely-available Websites. Among the next top five sources on the IU top 20 list (Table 2: WorldCat, Dissertation Abstracts, Lexis-Nexis Academic, Web of Science, Google Scholar), only the last is freely-available. Among the next top five sources on the QP top 20 list (Table 3: Google, Wikipedia, Answers.com, EBSCO, About.com), only one is fee-based. Google is ranked among the top sources on both lists, but it is Google Scholar at IU and Google.com on QP. Wikipedia is also on both top 20 lists: 10th place for IU and number 4 on the QP list. Ruffner and Abels (2005) reported a similar difference between academic and public libraries' use of free and fee-based sources in virtual reference.

Comparing Bradford et al.'s (2005) analysis of sources used to answer questions at the reference desk at an academic library with the IU transactions reveals that the top sources on both lists are the library catalog (15.86% for Bradford at al.; 20% IU), library Web pages (12.8% Bradford et al.; 30% IU), and database and indexes (25.2% Bradford et al.; 19% IU) (Tables 2 and 6). The variations between the two might be explained by 1) the use of different categories for analysis in the two studies; 2) different methods of counting sources used (librarian reports in Bradford et al.'s study or transactions transcripts in the present study); 3) variations in institutional practices or user needs between the two libraries; or 4) reflection of the continuing evolution of Web-based sources of information over time. The similarities are especially noticeable in the concentration of use on a small number of resources. The top five categories (out of 21) in Bradford et al.'s study accounted for 88.2% of the answers; in the current study, the top 20 sources (out of 850 titles) of the IU transactions accounted for 75% of the titles in the answers. The distributions in both studies are power curves, represented by the equation $y=ax^b$ where $a>0$. This is an extremely skewed distribution first mentioned in connection with acknowledgments (Davis & Cronin, 1993). It is typical in bibliometric analyses.

The librarians answering the QP transactions mentioned sources more frequently than did the IU librarians (53% to 47% of their answers; see Table 1). The QP librarians gave URLs in their responses more than four times out of five, compared with about half the time for the IU librarians (84% vs. 56%). The QP librarians also mentioned more sources per answer (1.6, compared with 1.4 average for the IU librarians). These differences in practice likely reflect the different kinds of questions being handled: 58% of known-item questions cited

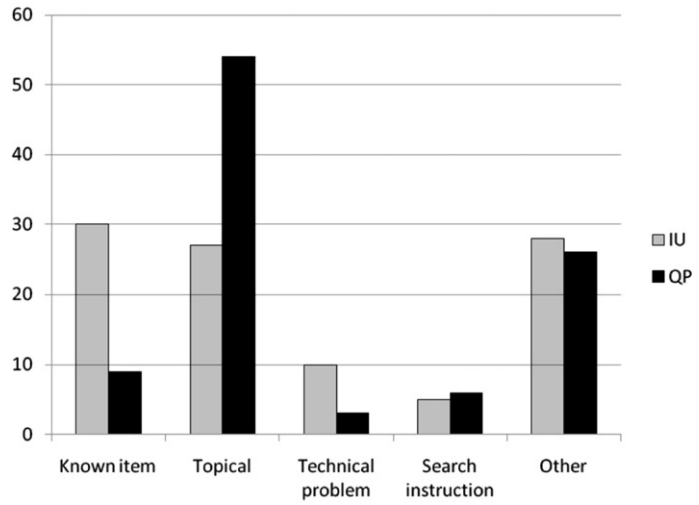

Fig. 1. Percentage of questions by type.

only one source, but only 31% of topical questions were answered as specifically. The most frequent types of questions were topical and known-item; together they accounted for more than half the questions in each data set. However, the numbers are roughly equivalent for the IU set (30% known item, 27% topical) but much more heavily skewed toward topical questions in the QP transactions (9% to 54%). Fig. 1 shows the distribution of question types.

Researchers compared the distribution of sources among the various LC classifications with Bradford's (2005) study of reference use of print sources at an academic library (Table 5). Fig. 2 shows the distribution by classification number. The profiles are remarkably similar, with heavy use in the social sciences (H) and literature (P). Bradford reported higher use than IU and QP in these areas and lower use in the general works (A), technology (T), and bibliography/library science (Z) classes. Many of the Web-based, uncataloged sources would likely be general works, technology, or bibliography, too. This would increase the disparity between Bradford's print-based distribution and the sources used for virtual reference. Without giving actual numbers, Welch, Cauble, and Little (1997) reported heaviest print reference collection use in the encyclopedias (AE), business (HF–HG), education (L, LB), literary criticism (PN), and law (KF) in their academic library. Their findings suggest that Bradford's low levels of use of generalities may reflect an anomaly in that library.

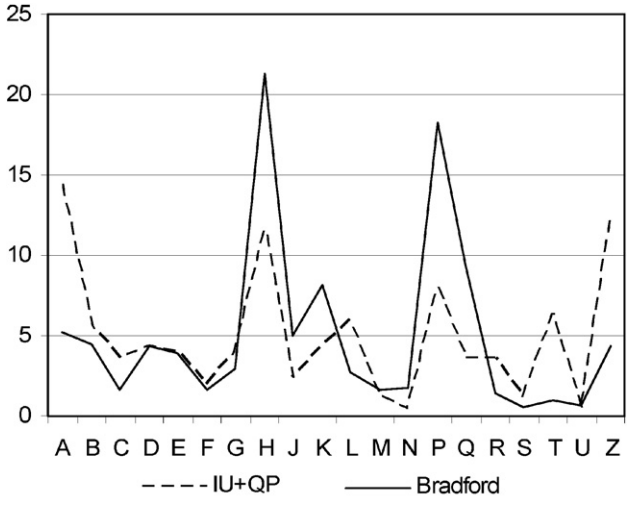

Fig. 2. Frequencies of use by LC classification.

6. Conclusion

Reference librarians today face the continuing growth of virtual reference, coupled with rising expectations for accountability in allocation and use of human expertise. Consortial or collaborative reference services offer one way to extend services; automated question answering or human support for automated systems have also been proposed (American Library Association, 2007; Pomerantz, 2005). Better understanding of how virtual reference interactions are conducted will improve the chances for success in any of these new applications.

This study supports the observation that librarians answering questions virtually rely on a concentrated set of sources of information–a core virtual reference collection–with a skewed bibliometric distribution. Almost all sources used (96%) were electronic; academic libraries tended to make greater use of fee-based sources but public libraries more often used sources freely-available on the Web. The increase in use of electronic sources by reference librarians should be taken into consideration for collection development decisions. E-mail reference questions to an academic library were nearly evenly divided between known-item and topical questions. Chat questions to both academic and public libraries were much more likely to be topical, and librarians answering these question mentioned more sources.

Additional studies are needed to test whether the trends identified here are evident in other libraries and continue over time. Studies comparing virtual with in-person reference transactions conducted by the same reference librarians, at the same reference desk, and during the same period would be especially useful in assessing the apparent increasing reliance on electronic sources. Comparisons of academic and public libraries' services could improve understanding of the differences in types of questions asked and preferences for fee-based or freely-available resources in both types of libraries. This study and future knowledge will aid in developing guidelines for practice, preparing the librarians who will provide virtual reference service, and managing the collections that will support their work.

Acknowledgments

The authors greatly appreciate cooperation from the Indiana University Bloomington Libraries Reference Department in providing access to transcripts of e-mail reference interactions. Sarah Horowitz provided remarkably insightful and efficient support in preparing and

coding these data. OCLC, Inc. and the QuestionPoint 24/7 Reference Service generously provided transcripts of chat reference sessions.

Appendix A. Coding scheme

Code		
Source name and call number		Title given in reply
		Full citation
		LC call number
Fee		
Number of sources per message		Range
		Average
Electronic source		
URL provided		
Reference sources	Catalog	Local catalog
		WorldCat
		Other
	Other reference sources	Databases and indexes
		Census
		Encyclopedia
		Library directory/ Web pages/ services/policies
		University URL
	Non-reference sources	Book
		Journal
		Newspaper
		Government website
		International government website
		Outside URL
		Other
Type of question		Known item
		Topical question
		Technical problem
		Search instructions
		Citing instructions
		Verification
		Non-reference
		Other
Type of library		Academic
		Public
		Unclear

References

Allen, B., & Reser, D. (1990). Content analysis in library and information science research. *Library & Information Science Research*, *12*, 251–262.

American Library Association. (2007). Indiana U, ChaCha partner on search platform. *American Libraries*, *38*(9), 42.

American Library Association, Reference and User Services Association. (2004). *Guidelines for behavioral performance of reference and information service providers*. Retrieved April 24, 2007, from http://www.ala.org/ala/rusa/rusaprotools/referenceguide/guidelinesbehavioral.htm

Berkov, E., & Morganstern, B. (1990). Getting to the core: Training librarians in basic reference tools. *Reference Librarian*, *30*, 191–205.

Bolander, R. C., Connaway, L. S., & Radford, M. L. (2006). *Seeking synchronicity: OCLC, Rutgers researchers explore virtual reference services by analyzing chat transcripts. NextSpace: The OCLC Newsletter, 1*(2) Retrieved April 16, 2007, from http://www.oclc.org/nextspace/002/research.htm

Bradford, J. T. (2005). What's coming off the shelves? A reference use study analyzing print reference sources used in a university library. *Journal of Academic Librarianship*, *31*, 546–558.

Bradford, J. T., Costello, B., & Lenholt, R. (2005). Reference service in the digital age: An analysis of sources used to answer reference questions. *Journal of Academic Librarianship*, *31*, 263–272.

Bradford, S. C. (1950). *Documentation.* Washington, DC: Public Affairs Press.

Carlson, S. (2007). Are reference desks dying out? *Chronicle of Higher Education*, *53*(33), A37.

Colson, J. (2007). Determining use of an academic library reference collection: Report of a study. *Reference & User Services Quarterly*, *47*, 168–175.

Davis, C., & Cronin, B. (1993). Acknowledgments and intellectual indebtedness: A bibliometric conjecture. *Journal of the American Society for Information Science*, *44*, 590–592.

Gotsick, P. S., Friedman, R. L., & Smith, J. B. (1979). The development and evaluation of a small ready-reference library collection for a rural practice: A case study. *Bulletin of the Medical Library Association*, *67*(2), 218–225.

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

Indiana University Libraries. (2008). About the Ask A Librarian e-mail service. Retrieved March 29, 2008, from. http://www.indiana.edu/~librcsd/ask/aboutask.html

Keenan, S., & Johnston, C. (2000). *Concise dictionary of library and information science*, (2nd ed.). London: Bowker Saur.

Kovacs, D. K. (2007). *Virtual reference handbook: Interview and information delivery techniques for the chat and e-mail environments.* New York: Neal Schuman.

LISWiki. (2007). *Chat reference libraries*. Retrieved April 5, 2007, from http://liswiki.org/w/index.php?title=Chat_reference_libraries&oldid=12730

McClure, C. R., Lankes, R. D., Gross, M., & Choltco-Devlin, B. (2002). *Statistics, measures and quality standards for assessing digital reference library services: Guidelines and procedures.*Syracuse, NY: Information Institute of Syracuse. Retrieved April 8, 2007, from http://quartz.syr.edu/quality/

Miles, M. B., & Huberman, A. M. (1994). *Quantitative data analysis: An expanded sourcebook*, (2nd ed.). Thousand Oaks, CA: Sage.

Mizzy, D., & Mahoney, E. T. (2002). Stocking the virtual reference collection. *Reference Librarian*, *38*(79/80), 67–99.

Nichols, M. I. (1993). *Selecting and using a core-reference collection*, (2nd ed.). Austin, TX: Texas State Library, Department of Library Development (ERIC Document ED377844).

OCLC, Inc. (2007). *Waterland Public Library, the Netherlands, fields 3 millionth question in QuestionPoint*. Retrieved March 29, 2008, from http://www.oclc.org/news/briefs/brief172.htm

O'Leary, M. (2003). QuestionPoint fortifies libraries in Internet age. *Online*, *27*(3), 70–72.

Pomerantz, J. (2005). A conceptual framework and open research questions for chat-based reference service. *Journal of the American Society for Information Science and Technology*, *56*, 1288–1302.

Ruffner, M., & Ables, E. (2005, November). *Digital answers: Connecting resources & responses*. Paper presented at the Virtual Reference Desk Conference, San Francisco, CA. Retrieved January 28, 2008, from http://ella.slis.indiana.edu/~shachaf/S551Spring2008/schedule.html

Shachaf, P. (2007). Virtual reference services: Implementation of professional and ethical standards. *Bulletin of the American Society for Information Science and Technology*, *34*(2), 20–24.

Shachaf, P., Meho, L. I., & Hara, N. (2007). Cross-cultural analysis of e-mail reference. *Journal of Academic Librarianship*, *33*, 243–253.

Smith, L. C. (2001). Selection and evaluation of reference sources. In R. E. Bopp, & L. C. Smith (Eds.), *Reference and information services: An introduction* (pp. 309–330). Englewood, CO: Libraries Unlimited.

Tenopir, C., & Ennis, L. (2002). A decade of digital reference: 1991–2001. *Reference & User Services Quarterly*, *41*, 264–273.

Virtual Reference Desk Network. (2003). *Facets of quality for digital reference services*. Retrieved August 2, 2007, from http://webjunction.org/do/DisplayContent?id= 11811

Virtual Reference Desk Project. (2003). *Facets of quality for digital reference services* (Version 5). Retrieved April 8, 2007, from. http://www.vrd.org/facets-06–03.shtml

Welch, J. M., Cauble, L. A., & Little, L. B. (1997). Automated reshelving statistics as a tool in reference collection management. *Reference Services Review*, *79*(3/4), 79–85.

White, M. D. (2001). Digital reference services: Framework for analysis and evaluation. *Library and Information Science Research*, *23*, 211–231.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology.* Cambridge, MA: Addison-Wesley.