Testing the Viability of the READ Scale (Reference Effort Assessment Data)©: Qualitative Statistics for Academic Reference Services

Bella Karr Gerlich and G. Lynn Berard

The READ Scale (Reference Effort Assessment Data) is a six-point scale tool for recording qualitative statistics by placing an emphasis on recording effort, knowledge, skills, and teaching used by staff during a reference transaction. Institutional research grants enabled the authors to conduct a national study of the READ Scale at 14 diverse academic libraries in spring of 2007 and test its viability as a tool for recording reference statistics. The study data were collected from 170 individuals and 24 service points with over 22,000 transactions analyzed. There was a 52 percent return rate of an online survey of participants, with more than 80 percent of respondents indicating they would recommend or adopt the Scale for recording reference transactions. The authors suggest that the READ Scale has the potential to transform how reference statistics are gathered, interpreted, and valued. This paper presents the findings of a nationwide study testing the Scale in spring 2007 and suggests practical approaches for using READ Scale data.

eference transactions are on the decline, as documented by librarians and their institutions, yet reference activities taking place beyond traditional service desks are on the rise. Librarians are reporting that they are as busy as they have ever been. According to an Association of Research Libraries (ARL) 2002 study conducted to reveal best practices in reference work, the findings exposed a general lack of confidence in current data collection techniques as "failing to capture and accurately reflect reference activities overall."[1]

What factors account for this change in reference work? Technology has transformed our ability as information providers to serve our user communities, structure our facilities, and conduct

Bella Karr Gerlich is University Librarian in the Rebecca Crown Library at Dominican University; e-mail: bkarrgerlich@dom.edu. G. Lynn Berard is Principal Librarian, Engineering and Science, and Fellow of the Special Libraries Association in the University Libraries at Carnegie Mellon University; e-mail: lberard@andrew.cmu.edu. This paper was presented at the RUSA RSS Research & Statistics Committee's 14th Annual New Reference Research Forum at the 2008 ALA Annual Conference in Anaheim, California, on Sunday, June 29, 2008. Select portions (such as the READ Scale©, concept, and figures) of this paper have appeared in other publications in shorter, focused, introductory articles. This paper comprises the complete 2007 national study results. The READ Scale (Reference Effort Assessment Data) ©Bella Karr Gerlich

our work. The introduction of online information resources has heightened the need for instruction in the classroom, as well as instruction via e-mail, over chat services, and at point of use. Reference librarians are being sought out for their knowledge management expertise and subject specialization at the reference desk as well as increasingly in their offices and hallways. Counting traffic numbers at the traditional reference desk is no longer sufficient as a measurement that reflects the effort, skill, and knowledge associated with this work.

FIGURE 1
READ Scale—Reference Effort Assessment Data Scale©

1: Answers that require the least amount of effort and no specialized knowledge skills or expertise. Typically, answers can be given with no consultation of resources. Length of time needed to answer these questions would be less than 5 minutes. Examples: directional inquiries, library or service hours, service point locations, rudimentary machine assistance (locating or using copiers, how to print a document or supplying paper).

2: Answers given that require more effort than the first category but require only minimal specific knowledge skills or expertise. Answers may need nominal resource consultation. Examples: call number inquiries, item location, minor machine and computer equipment assistance, general library or policy information (how to save to a disk or e-mail records, launching programs or rebooting).

3: Answers in this category require some effort and time. Consultation of ready reference resource materials is needed; minimal instruction of the user may be required. Reference knowledge and skills come into play. Examples: answers that require specific reference resources (encyclopedias or databases); basic instruction on searching the online catalog; direction to relevant subject databases; introduction to Web searching for a certain item; how to scan and save images; more complex technical problems (assistance with remote use).

4: In this category, answers or research requests require the consultation of multiple resources. Subject specialists may need to be consulted and more thorough instruction and assistance occurs. Reference knowledge and skills needed. Efforts can be more supportive in nature for the user or, if searching for a finite answer, difficult to find. Exchanges can be more instruction based as staffs teach users more in-depth research skills. Examples: instructing users how to use complex search techniques for the online catalog, databases, and the Web; how to cross-reference resources and track-related supporting materials; services outside of reference become utilized (ILL, Tech services, etc.), collegial consultation; assisting users in focusing or broadening searches (helping to redefine or clarify a topic).

5: More substantial effort and time spent assisting with research and finding information. On the high end of the scale, subject specialists need to be consulted. Consultation appointments with individuals might be scheduled. Efforts are cooperative in nature, between the user and librarian and/or working with colleagues. Multiple resources used. Research, reference knowledge and skills needed. Dialogue between the user and librarian may take on a "back and forth question" dimension. Examples: false leads, interdisciplinary consultations/research; question evolution; expanding searches/resources beyond those locally available; graduate research; difficult outreach problems (access issues that need to be investigated).

6: The most effort and time expended. Inquiries or requests for information can't be answered on the spot. At this level, staff may be providing in-depth research and services for specific needs of the clients. This category covers "special library" type research services. Primary (original documents) and secondary resource materials may be used. Examples: creating bibliographies and bibliographic education; in-depth faculty and Ph.D. student research; relaying specific answers and supplying supporting materials for publication, exhibits etc; working with outside vendors; collaboration and ongoing research.

Gerlich developed the READ Scale at Carnegie Mellon University as a proposed quantitative measurement method designed to capture all occurrences of reference activity.[2] The READ Scale (Reference Effort Assessment Data) is a six-point scale used for recording vital supplemental qualitative statistics gathered when reference librarians assist users with their inquiries or research-related activities by placing an emphasis on recording the skills, knowledge, techniques, and tools used by the librarian during a reference transaction (figure 1).

Institutional grants received in 2006 enabled the authors to expand the study beyond one institution to fifteen academic libraries in the spring of 2007 with the goal of testing the viability of the READ Scale as an adaptable tool for gathering qualitative statistical reference data on a national level.

Study Objective

Our objective was to test the viability of the READ Scale as an additional tool for gathering reference statistics. The READ Scale was launched at Carnegie Mellon University as a trial in the spring of 2003, followed by an academic year study in 2003–2004. The READ Scale emphasizes effort, skills used by staff at the time the reference transaction occurs. This method is especially appealing in a profession where the current industry standard for recording statistical data is a hash mark that records and recognizes quantity as opposed to quality.

Literature Review

A review of literature and studies on reference librarians, reference services, and reference statistics was used to inform and support the design of the READ Scale, as well as the contribution of qualitative study to librarianship. There are two distinct areas of study in reference assessment that directly influence our work: the measurement and evaluation of reference service and the means of recording reference transactions (both traditional and automated practices).

Measurement and Evaluation of Reference Service

Beyond efficacy; the exemplar librarian as a new approach to reference evaluation by Quinn (1994)[3] takes an interesting approach as it suggests using qualitative methods of evaluating reference librarians by first asking "what makes a reference librarian great?" Quinn asserts his study implies that good reference behavior is learned and that cultural preparation is a must. The study also found that not one single factor made a librarian great: it is a combination of skills. Quinn's article focuses on behavioral aspects of reference librarianship. This study will add to those findings by determining if participants' using the READ Scale find that ranking and recording their efforts results in positive feelings as their effort, skills, and knowledge are being recognized during the reference transaction.

Quality Reference Service: A Preliminary Case Study, Stalker and Murfin (1996)[4] studied the results of the WOREP (Wisconsin-Ohio Reference Evaluation Program) survey at Brandeis University, which demonstrated the highest level score to date of a general reference department using the WOREP, to determine to what extent the high quality of professional service was demonstrated, due to use of the WOREP model. This article found that allowing for sufficient time for the consulting role of reference librarians led to the high success rate when using the WOREP Model at Brandeis; other factors included contents and configuration of the reference area, and strong support for services by administration. The READ Scale likewise acknowledges the interactive nature of the reference transaction, the time element and records the service component.

Perspectives on Quality of Reference Service in an Academic Library: A Qualitative Study was a study done by Mendelsohn (1997)[5] to explore the concept of quality as it applies to reference service. Four participants in humanities and social sciences areas were interviewed and perceptions of quality discussed. This paper supports earlier works that emphasize willingness

to help, knowledge and skills, morale and time as vital components in the quality of the reference transaction from the librarian point of view.

Work in Motion / Assessment at Rest: An Attitudinal Study of Academic Reference Librarians; A Case Study at Mid-Size University MSU A, written by Gerlich (2006)[6] is a study that focuses solely on reference librarians and their attitudes about their work: what they value, how they perceive themselves, how they perceive others view them. This study supports the notion that reference, or the transaction interaction, is the primary function of the reference librarians' position and the highest valued task by both the reference librarians and administrators. The study also reveals a lack of assessment or reward for this work outside of the anecdotal, with librarians and administrators in agreement that current statistical data gathered for reference work is not adequate for recording effort, knowledge, and skill.

Testing Classification Systems for Reference Questions, Henry and Neville (2008)[7] follows the University of South Florida, St. Petersburg study using Warner's classification system at the Nelson Poynter Memorial Library in comparison to Katz's traditional reference categories described in Introduction to Reference Work (directional, ready reference, specific search questions and research). The results of this study support the idea that the adoption of new measures for reference statistics seems warranted to be more exacting, relevant, and reflective of reference services. The conclusions also reached similar findings of the READ Scale that by recording actual effort means reexamining staffing of the reference desk as a service point.

The Recording of Reference Transactions

Usage-Based Staffing of the Reference Desk: A Statistical Approach, Dennison (1999)[8] discussed the importance of staffing decisions for reference desks, and how measuring usage of service can inform those decisions. At Winona State University Library (WSU), Dennison reports on using direct measurement applications to reference statistics. WSU employed categories to record reference statistics and determine peak times for staffing the reference desk based on the category assigned to each transaction.

A New Classification for Reference Statistics by Warner (2001)[9] describes a test of an alternative reference data-gathering model. The impetus for creating the classification model in Warner's case was borne out of need for training and triage at a new single point-of-service desk at Eastern Carolina University. Warner's study changed from a daily collection of data for the first three months to being randomly selected once a month. Warner's research and subsequent implementation of a classification system in this case lays a foundation for this study by introducing alternative methods for gathering statistics.

SPEC Kit 268, Reference Service Statistics & Assessment, Novotny (2002)[10] paints a picture of changing reference services and stagnant assessment measures of the same in research libraries by surveying and documenting how ARL libraries were collecting and using reference service transactions data. This survey described in its executive summary the confusion and angst surrounding modern reference work as libraries scramble to collect data. There is no mention of improving reference quality, developing employees, or recognition of work effort—the study did not distinguish between a successful or unsatisfactory transaction. While it recognizes the use of electronic tools to gather data, there is a failure to recognize the librarian's use of electronic tools to distribute information in any sense outside the narrow confines of the "transaction" definition. This study was most useful for this work in that it painted a picture that the system of reference assessment in use by ARL libraries appears to be in flux.

Reference Use Statistics: Statistical Sampling Method Works (University if Tennessee at Chattanooga) by Murgai (2006)[11] supports one of the findings of the Novotny study

that librarians felt busier than ever helping patrons, despite a decline in the number of patrons served. Murgai suggested that most reference librarians would like reference statistics to reflect all aspects of reference but would also like statistic recording to be simple, while acknowledging that reference service is anything but simple. The University of Tennessee at Chattanooga (UTC) reviewed other academic libraries' sampling methodologies and employed sampling for a year to compare to daily data gathering. The results of the statistical analysis showed that the numbers gathered for a set period of time are very close to data gathered over a longer period of time, supporting the results of the 3-week period of data capture selected for the READ Scale study. The limits of the UTC study also support the need for

a tool like the READ Scale, noting that the classifications for reference statistics used in the UTC study did not capture the types of questions, resources, off-desk questions—measures that are used in the READ Scale—were needed to get a complete picture of reference services.

Methodology
Timeline
The preparation of this study occurred in the summer and fall of 2006, with participation commitments in place by late November 2006. The Institutional Review Board (IRB) approval and pre-study exercises took place between December 1 and February 4.

Libraries were given the option of conducting the study for the duration of their spring semester, and or for the predeter-

TABLE 1		
READ Scale Participating Institutions		
Enrollment Less than 5,000–5 Institutions	**Enrollment Greater than 5,000–4 Institutions**	**Enrollment Greater than 10,000–5 Institutions**
Clarke College Clarke College Library Dubuque, IA	Carnegie Mellon University (1 Institution, 6 Service Points) Pittsburgh, PA	Georgia Institute of Technology Georgia Tech Library Atlanta, GA
Eastern Virginia Medical School Edward E. Brickell Medical Sciences Library Norfolk, VA	Georgia College & State University Library & Instructional Technology Center (1 Institution, 2 Service Points) Milledgeville, GA	New York University Business & Documents Center – Bobst Library New York, NY
Lawrence University Seeley G. Mudd Library Appleton, WI	Robert Morris University (1 Institution, 2 Service Points) Moon Township, PA	West Virginia University (1 Institution, 3 Libraries) Morgantown, WV
Lewis & Clark College Aubrey R. Watzek Library Portland, OR	Washburn University Mabee Library Topeka, KS	University of California, San Diego Science & Engineering Library La Jolla, CA
Our Lady of the Lake University San Antonio (OLLUSA) Sueltenfuss Library San Antonio, TX		University of Nebraska Love Library (Chat Service only) Lincoln, NE

mined three-week duration: February 4–February 24, 2007. These two options were selected to accommodate those institutions that normally only sample reference statistics as well as those that collect data daily for an entire semester. All institutions had to commit to the February data collection period. These three weeks were selected to limit the chance for spring breaks to occur within the study time frame.

Study Participants

The research team decided on the following parameters for seeking participants in the study. The universities must:

• Number between 9 and 15 academic libraries

• Be diverse geographically

• Contain diverse enrollment figures, grouped as follows: ≤5,000, >5,000 and ≤10,000, and ≥10,000

• Include both public and private institutions

The number range 9–15 was determined with a minimum acceptance rate of 9 participating, with at least three for each enrollment figure represented. One institution that initially agreed to participate had to withdraw for reasons unique to that university, leaving the number of participants at 14 institutions, with 170 individual participants total. See table 1 for participating libraries. Each institution was asked to identify an onsite coordinator at each location who would commit to disseminating information and managing the activities, timelines and follow-up associated with conducting the study.

Pre-Study Calibration of Sample Scale Questions

To familiarize and prepare participating librarians with the READ Scale and its proper use, a list of pre-study test questions was developed and sent to onsite coordinators. Each site received the same set of questions; however, the coordinators were instructed to select some questions from the list but were given the flexibility to substitute others for those localized to the institution. The addition of a sample

question(s) that occurs frequently at the home institution reference desk provided a common ground for a discussion of how to apply the scale when rating the effort level of the transaction. The least number of questions distributed was six, and we asked that a range of project effort be represented (1–6 levels on the READ Scale) to acquaint participants with the full range of scale levels. All participants were asked to answer and rank their effort for each of the sample questions. It was agreed that onsite coordinators would evaluate responses and respond to participants' questions regarding all aspects of applying the scale. Participants were also asked to record time during this exercise so that the researchers could average the length of time per transaction, per scale rank overall. Table 2 represents those questions from the researchers' test list along with the average time it took to complete the transaction.

Across the board, the pre-study rating effort for transactions at the 1, 2, or 6 level were typically unanimous, while the 3, 4, and 5 ratings revealed some differences between individuals' perceived rankings. Differing of individual rankings for the same type of reference transaction was thought to be due in part to subject specialization and how individuals tend to "grade" (hard or easy). Coordinators met with their participants and summed up how the transactions were resolved, the recommended rating to assign, the time it took to answer the question, and the reason for the rating. This enabled individuals to adjust their personal grading habits for traditional inquiries. It was important to recognize that where subject specialization is the norm, effort associated with customer service should be recognized. This is why the number of elements (the definition for each number on the scale) and time associated with the scale rankings are important to note. Staff helping someone out of their area of expertise should feel comfortable assigning a higher scale point than the librarian with a specialization in the subject area. As noted later in this paper, the criteria of time and

how it is applied using the READ Scale is an area considered for further research.

Additionally, reference librarians were asked to conduct the study in their offices during *"off-desk"* times. The term *"off-desk"* is used to note reference transactions handled by a reference professional that occurs away from an established, regularly

TABLE 2 Common Test Questions	
Common Question, Academic Libraries	**Most Common Ranks & Average Time**
I need a translation for an Italian aria.	3, 7 min
Where is the bathroom?	1, 1 min
I have a laptop—where I print out library records?	2, 1 min
I am researching postwar suburban housing development in the (City) region—can you show me what you have in that relates to this topic, or where I should look?	5–6, 90 min
I need the issue number for this citation: Le Goff, Jacques Ordres mendiants et urbanisation dans la France Annales: *conomies, soci*t*s, civilisations, vol. 25, (1970)	4, 15 min
I am trying to place a hold on a book in process by using the online catalog request form. Kept receiving error message requesting item info—please help!	2, 5 min
I'm trying to find out about the philosophy of St. Benedict. Do you have any suggestions on which of his books or writing I can download?	4, 15 min
I need to find some contemporary criticisms for the play *Fences* by August Wilson—both the writing of the play and a production.	4, 12 min
I am looking for some help getting started on a research project —gender roles and the selection of college majors in the south— where do I start? How do I conduct a study?	5, 23 min
Do you have a book with pictures of kitchen utensils used in colonial times?	4, 28 min
PsycInfo says that we have this journal, but it isn't in the library —please help!	2–3, 5 min
Common Questions, Academic Medical Libraries	
Curriculum models for teaching medical students about medical ethics: (1) What should be the learning objectives; 2). What the curriculum content should entail.	5–6, 90 min
Need recent (up to 10 years) clinically relevant articles on the patient care of thrombolytic therapy and antiplatelet therapy and anticoagulation in the treatment of peripheral vascular disease	3–4, 15 min
I need a list of drugs that affect lymph flow or lymph vessel contractions.	3–4, 15 min
I'm looking for medical licensure lookup, medical school etc and if there are any malpractice proceedings against Dr. _____—can you help?	3, 10 min

TABLE 3 Sample Form with Categories and Approach Type											
Date:							Walk-Up Dir	Walk-Up Ref	Phone Dir	Phone Ref	E-mail
SCALE	1	2	3	4	5	6					
Walk-Up Directional											
Walk-Up Reference											
Phone Directional											
Phone Reference											
Chat											
E-mail											

scheduled reference desk. Anecdotal evidence suggests that this is where the majority of higher-level scale effort in assisting patrons is being conducted, especially for those clients served by a liaison librarian with subject-specific responsibilities. These data were gathered and compiled to help determine at which service point users sought assistance; it was theorized that transactions at the 4, 5, and 6 levels would be recorded by individuals while working from their offices rather than at a traditional service point. The recording of *"off-desk"* statistics is a nontraditional activity and one not often employed by reference librarians or reported institutionally. It is the case then that this valuable effort has not been seriously studied or credited to the work effort of reference professionals.

The READ Scale data recording method is such that it allows institutions to use their local paper or online form that captures day, hour, and approach type for both directional and reference questions, on and off desk. Participants in the study were asked specifically to use their existing forms to test the adaptability or translation of the READ Scale in using a number from the scale in place of a hash mark when recording a reference transaction. On the researchers' end, there was little difficulty in recording data onto the statistics spreadsheets, and the benefit for participating institutions was the ease of adoption of the scale into existing local recording instruments.

Data Collection

As all of the institutions had different methodologies in place for recording statistics, researchers developed a common table to compile data by Scale number and approach type (table 3).

Some institutions had numerous categories that identified inquiry types, such as "equipment" or "database search." These were placed into the "Walk-Up Reference" category for the study. READ Scale definitions do not distinguish the kind of question, but they reflect the effort expanded, knowledge required, and even the teachable moment that occurs during the transaction.

The time of day that the transaction occurred was not reported cumulatively by the researchers, as reference desk hours and personnel schedules varied by institution and could not be normalized. They were recorded for each individual institution as reported and made available to the respective organization so that assessments could be made locally.

The approach type for transactions was recorded to establish frequencies for how transactions occurred. As suggested by the ARL study, some academic institutions are experiencing a decline in reference transactions. Recording approach frequency here would help determine the most popular method for seeking reference help and where that transaction occurred.

At the conclusion of the three-week data collection period, an online question-

naire was sent to all study participants. The survey was designed to assess the participants' experience when applying the scale, to gain their feedback on the value of the scale in demonstrating effort when recording reference transactions via this method, and to inquire how the scale might be changed to improve the data collection instrument. While the researchers' individual institutional experiences with the scale were very positive, one desired outcome for conducting a national study was to determine the viability of the 1–6 point Scale.

Results

Three-Week Study

Fourteen institutions participated in the READ Scale Study during the spring semester of 2007. There were a total of 24 service points and 170 individual participants. All institutions submitted statistics using the READ Scale for the same three-week time period, February 4–February 24, 2007. Seven institutions elected to continue using the Scale for the duration of their respective semester *after* the initial study period. Table 4 illustrates the cumulative number of transactions, READ Scale category assignment, question and approach type for all service points, and institutions for a total of 8,439 transactions during the three-week study period. All institutions were encouraged to use the READ Scale for recording off-desk statistics as well, if appropriate. Seventeen out of a possible 170 individuals reported off-desk statistics for a total of 1,531 off-desk transactions recorded in the three-week period (table 5). Combined transactions for service points and off-desk totaled 9,970.

The study illustrated that the majority of inquiries continue to be by physical approach (figure 2). "Off-desk" the percentage of e-mail is considerably higher (figure 3) and almost equal in percentage to in-person interactions.
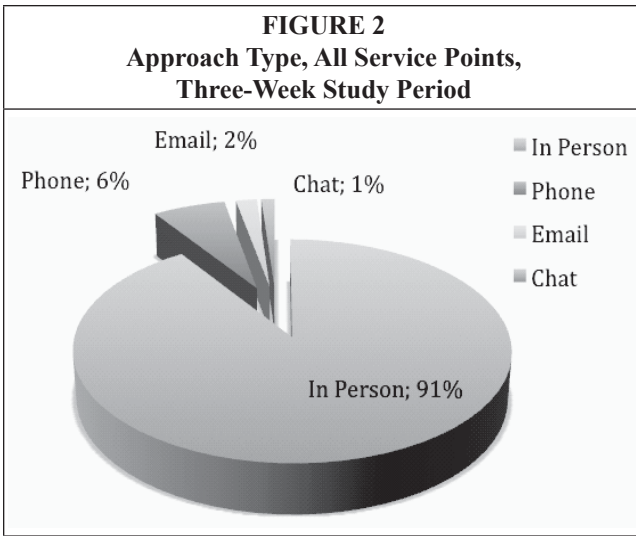
Comparisons between service points illustrate that the highest majority of

TABLE 4								
Cumulative Data, All Service Points, All Institutions, 2/4–2/24/07								
READ SCALE	1	2	3	4	5	6		
Walk-Up Directional	2,260	337	23	2	0	0		
Walk-Up Reference	1,693	1,750	1,067	397	89	34		
Phone Directional	148	38	4	5	0	0		
Phone Reference	111	113	85	17	7	5		
E-mail	47	44	44	19	0	2		
Chat	13	19	44	22	0	0		
Totals	4,272	2,301	1,267	462	96	41	Total	8,439

TABLE 5								
Cumulative Data, Off-Desk, All Institutions, 2/4–2/24/07								
READ SCALE	1	2	3	4	5	6		
Walk-Up Directional	23	4	0	3	1	0		
Walk-Up Reference	196	197	157	74	41	18		
Phone Directional	44	6	2	0	0	0		
Phone Reference	85	109	41	20	5	2		
E-mail	193	142	93	49	21	5		
Totals	541	458	293	146	68	25	Total	1,531

**FIGURE 2
Approach Type, All Service Points,
Three-Week Study Period**

Email; 2%

Phone; 6%   Chat; 1%

■ In Person
■ Phone
■ Email
■ Chat

In Person; 91%

of the higher-level effort, knowledge, and skill required of reference personnel will take place away from the public service point. The need to increase efforts to record off-desk reference statistics was also expressed by the many of the respondents in the ARL Study.[12]

Semester-Long Study
Seven of the institutions elected to continue to use the READ Scale for the duration of their respective semesters. Figures that follow represent fourteen service points and ninety-four individual participants. There were a total of 15,194 transactions recorded (table 6). Data was collected through May 11 and includes three-week study figures reported previously. Approach type for this group was also recorded (figure 6).

All institutions were encouraged to continue to use the READ Scale for recording off-desk statistics as well, if appropriate. Seven institutions, eight service points, and a possible 66 individuals reported off-desk statistics, for a total of 1,156 transactions recorded for the duration of their respective semesters (table 7). Data includes three-week study figures
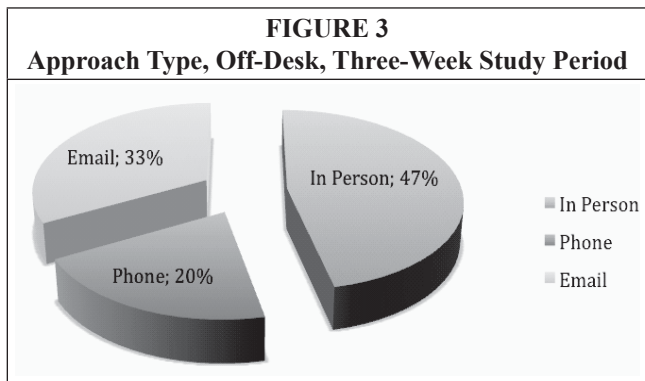
transactions that occur at the reference desks are in the READ Scale number-one category (figure 4), followed by number category two revealing that most inquiries at the public service point require the least amount of effort, knowledge, and skills of library personnel.

Off-desk comparisons show a different but consistent pattern (figure 5): that the percentage of questions answered off-desk for most of the institutions require a much higher level of effort, knowledge, and skills from reference personnel than at the public service point. Only three of the seventeen off-desk comparators in figure 5 have more level READ Scale categories representing the bulk of their transactions off-desk, with only two of those recorded in the higher than 40 percent range, overall. The majority of the off-desk ratings for the remaining group were at category two, three, and four respectively, suggesting that users actively seek out the expertise of particular reference staff.
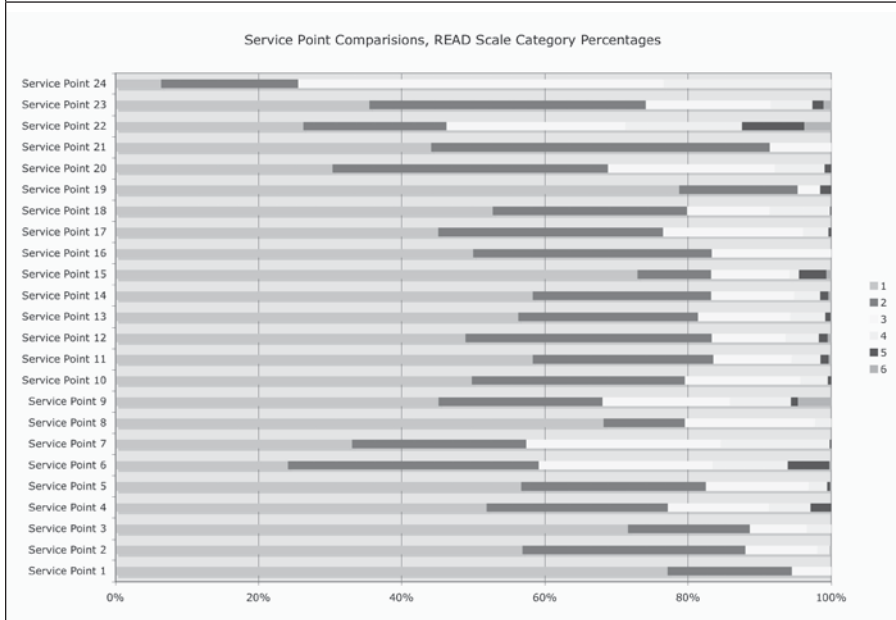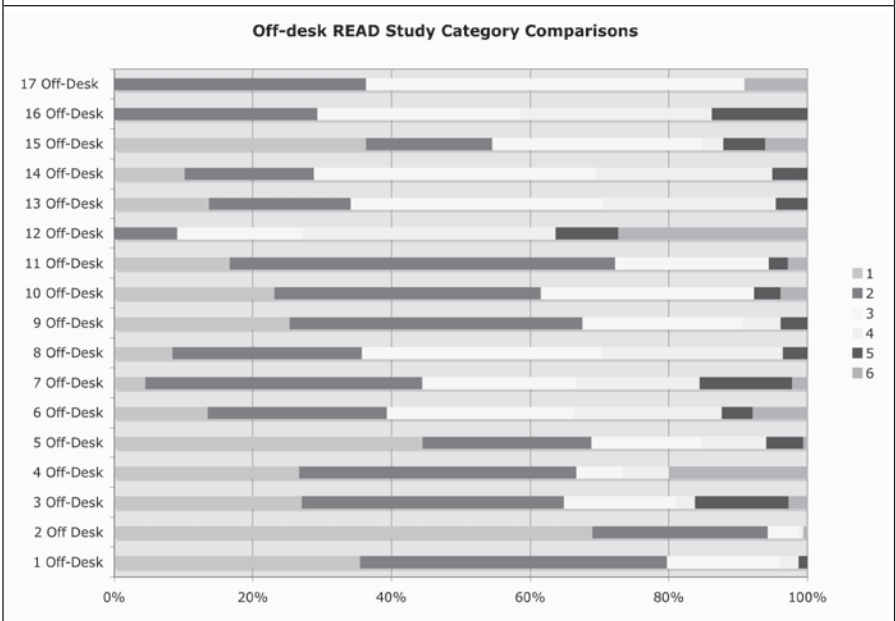
These data further support the researchers' theory that most

**FIGURE 3
Approach Type, Off-Desk, Three-Week Study Period**

Email; 33%

In Person; 47%

Phone; 20%

■ In Person
■ Phone
■ Email

FIGURE 4
Comparative Illustration of the Percentage of Each READ Scale Category,
per Service Point

FIGURE 5
Comparative Illustration of the Percentage of Each READ Scale Category,
Off-Desk

READ SCALE	1	2	3	4	5	6		
Walk-Up Directional	3,787	899	56	6	0	0		
Walk-Up Reference	2,203	2,606	1,784	501	153	29		
Phone Directional	377	148	10	4	1	0		
Phone Reference	375	358	231	40	22	3		
E-mail	465	423	238	85	69	4		
Chat	19	76	150	66	6	0		
Totals	7,226	4,510	2,469	702	251	36	Total	15,194

TABLE 6
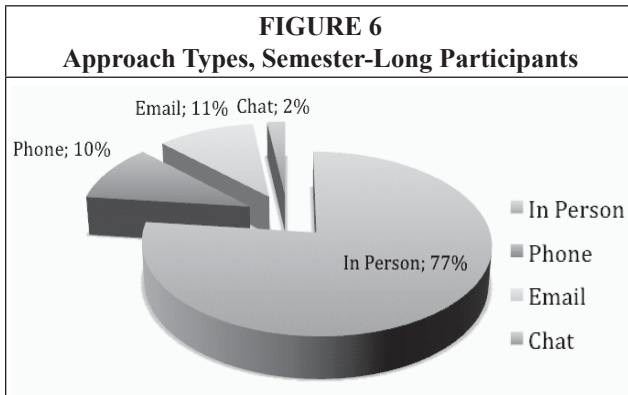**Cumulative Data, All Service Points, Semester-Long Participants**

reported previously. Approach type for transactions that occurred off-desk was also recorded (figure 7).

As with the three-week study period, the semester-long group's data compilation showed that the preferred approach type was "in person" overall. However, when separated out, the use of e-mail as an approach type came very close comparatively to that of in-person approach type when a transaction took place off-desk.

Comparative illustrations coincide with the three-week dataset; the majority of the transactions that occur are at category 1 of the READ Scale at service points for all institutions (figure 8).

Off-desk comparisons again show a different but consistent pattern (figure 9). The percentage of questions answered off-desk for the semester-long group participants required a much higher level of effort, knowledge, and skills from reference personnel than at the public service point. Unlike the three-week study, however, no scale category level 1 exceeded the 40 percent mark, and two of the off-desk institutions recorded no level 1 transactions at all. The semester-long off-desk group also

FIGURE 6
Approach Types, Semester-Long Participants

Email; 11% Chat; 2%
Phone; 10%
In Person; 77%
In Person
Phone
Email
Chat

READ SCALE	1	2	3	4	5	6		
Walk-Up Directional	2	1	1	1	0	0		
Walk-Up Reference	89	134	153	87	51	28		
Phone Directional	0	0	0	0	0	0		
Phone Reference	30	53	48	22	3	2		
E-mail	43	95	181	105	24	3		
Totals	164	283	383	215	78	33	Total	1,156

TABLE 7
Cumulative Off-Desk Data, Semester-Long Participants

FIGURE 7
Approach Types, Off-Desk,
Semester Long Participants
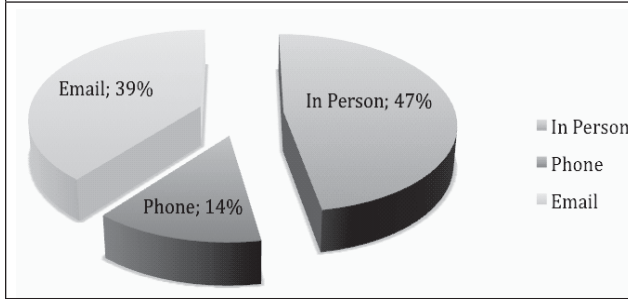
3 and 4 off-desk data suggest that as the semester continues, there is a likelihood that the opportunity for off-desk transactions increases and the need for a level of expertise, knowledge, and skill likewise increases. This coincides with curriculum expectations that typically have fewer difficult assignments at the onset of a semester but demonstrate an increase in complicated assignments and prolonged research projects as the term progresses.

reported a higher percentage of level 3 category transactions than 2 with level 4, 5, and 6 following.

As stated earlier, these data support the researchers' theory that most higher-level effort, knowledge, and skill required of reference personnel take place away from the public service point. Furthermore, increases in the percentages of READ Scale categories

Online Survey Results
An anonymous survey was constructed to solicit feedback on the READ Scale that included its ease of use, participant difficulty distinguishing between

FIGURE 8
Comparative Illustration of the Percentage of Each READ Scale Category,
per Service Point, Semester-Long Participants
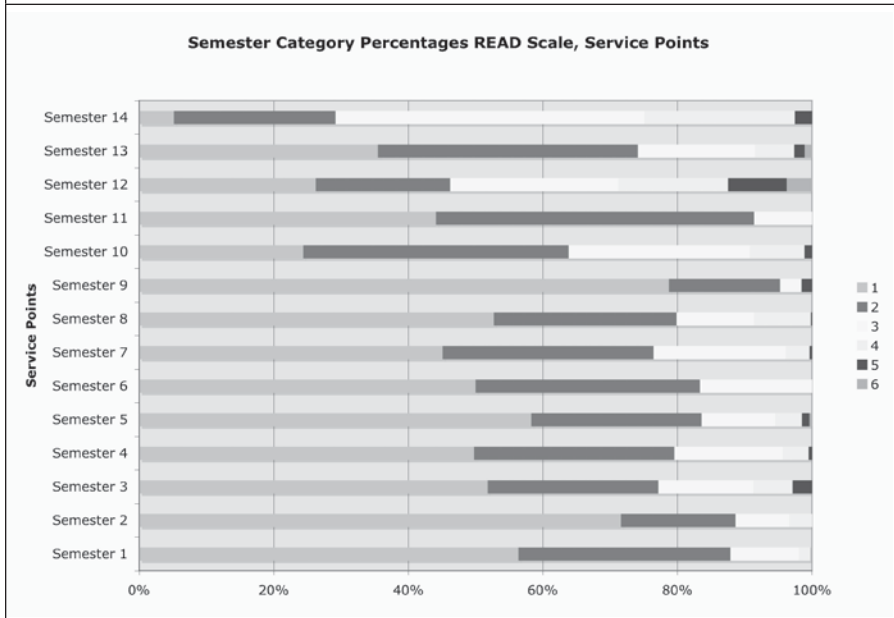
FIGURE 9
Comparative Illustration of the Percentage of Each READ Scale Category, Off-Desk, Semester-Long Participants

Off-Desk Semester Percentages
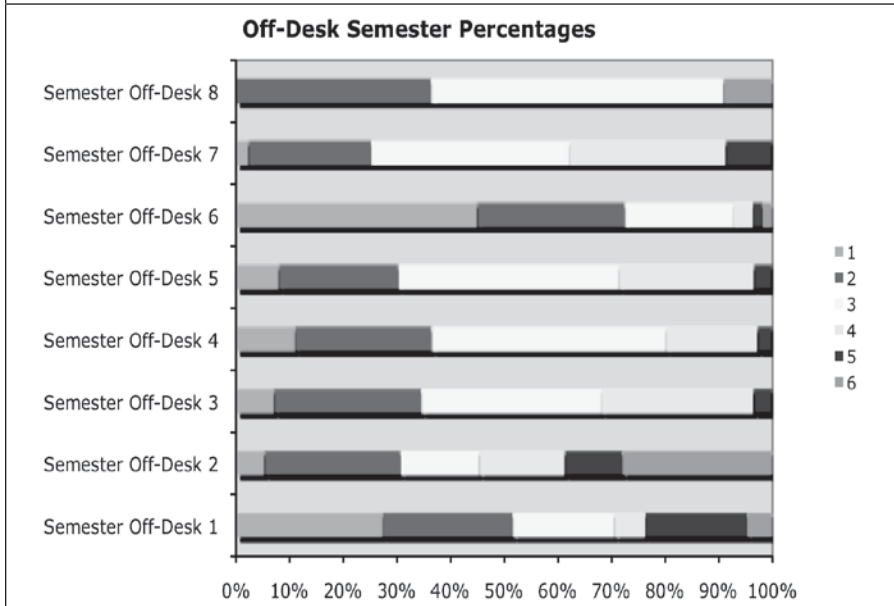
TABLE 8
Degree of Difficulty
Question: Please Rank Your Degree Of Difficulty Using The Read Scale

Responses	Not Difficult	Somewhat Difficult	Moderately Difficult	Difficult	Very Difficult	Skipped Question	Number Responded
	52 (51.0%)	38 (37.3%)	10 (9.8%)	2 (2.0%)	0	0	102

TABLE 9
Application Ease / Question: Was the READ Scale Easy to Apply?

Responses	Very Easy to Apply	Easy to Apply	Moderately Easy	Somewhat Easy	Not Easy	Skipped Question	Number Responded
	16 (15.7%)	39 (38.2%)	38 (37.3%)	8 (7.8%)	1 (1.00%)	0	102

TABLE 10
Scale Adds Value to Statistic Gathering
Question: Please rank the level of perceived "added value" the READ Scale placed on statistics gathering for reference transactions.

Responses	Extreme Value Added	High Value Added	Moderate Value Added	Minimal Value Added	No Value Added	Skipped Question	Number Responded
	7 (6.9%)	46 (45.5%)	35 (34.7%)	9 (8.9%)	4 (4.0%)	1 (.99%)	101

TABLE 11
Difficulty Between Rankings

Question: Did you have difficulty in deciding between ratings? If so check all that apply.

READ Scale	1-2	2-3	3-4	4-5	5-6	No Difficulty	Response Count	Total Responses	Skipped Question
Responses	12 (7.6%)	32 (20.4%)	46 (29.3%)	31 (19.7%)	15 (9.6%)	21 (13.4%)	157	99	3

Self-evaluation

Question: How did you feel about evaluating your own efforts?

Responses	Extremely Comfortable	Very Comfortable	Moderately Comfortable	Minimally Comfortable	Not Comfortable	Skipped Question	Number Responded
Responses	12 (11.9%)	50 (49.5%)	35 (34.7%)	4 (4.0%)	0 (0%)	1 (.99%)	101

categories, and participants' perception regarding added value to reference work. All participants (170) were sent an online survey to complete. The response rate for the survey was high, with 102 (60%) total respondents. The questions and their responses are detailed in this paper.

The majority of participants had no difficulty using the READ Scale (table 8) and found the READ Scale easy or moderately easy to apply (table 9). When asked to rank perceptions of added value to statistical data gathering, the majority of responses fell in the "high-value added" category (table 10). The favorable response rate, with the majority of respondents in agreement that the READ Scale's added value to reference statistics is "high" (45%) or "moderate" (35%) accounts for a total of 80 percent of the study group's opinions.

Participants were asked about difficulties they may have experienced in deciding between categories. Most implied difficulty deciding between ranks 3 and 4; participants were also asked how they felt about evaluating their own efforts (table 11) with the majority responding that they were comfortable with the process.

Asked if they would recommend the READ Scale to another reference librarian, 67 percent of the study participants answered in the affirmative as is, with another 20 percent who would recommend it with modifications, bringing the favorable response rate to more than 80 percent.

A follow-up question inquired if the study group would likely be in favor of having the Scale adopted in their library as is, or with modifications. A total of 50 percent responded affirmatively to "as is," with another 30 percent who would adopt with modifications, bringing the favorable response rate to 80 percent.

The survey group was also given an opportunity later in the survey to suggest modifications, of which 24 deposited comments and two optional questions asked for specifics about what the study group liked and disliked about the READ Scale.

The likes listed by the participants where coded into the six most common reoccurrences: Effort/Value; Approach to Evaluation; Types/Levels; Time; Staffing Levels; and Reporting to Administration:

Sample comment, Effort/Value (17 occurrences noted):
It gave me a quick visible check of my recent efforts. This made my deskwork more rewarding, since I sometimes feel like I do so many 1s and 2s—but I could see that I was actually doing a higher level of reference than I realized. It added value to the statistics—literally

Sample Comment, Approach to Evaluation (13 occurrences noted):
It qualifies what we were only quantifying and therefore is a more realistic indicator of what we do at the desk.

Sample Comment, Types/Levels (9 occurrences noted):
I like that it makes a qualitative distinction between types of reference interactions; it gives credit to more challenging transactions. The differences between the kinds of interactions are flattened in a typical "hash mark" approach to noting reference interactions.

Sample Comment, Time (5 occurrences noted):
I thought that it was a good way to see how the time was being spent on the question. It gives a better picture of what you are doing instead of just a tally mark for each question.

Sample Comment, Staffing Levels (6 occurrences noted):
Using the scale made me think about the types of questions we were receiving via the various formats and how we might need to change staffing patterns to better serve our users.

Sample Comment, Reporting to Administration (5 occurrences noted):
It will give a better contour to statistics as read by administrators and funders, and help to make better staffing decisions.

Dislikes were coded into the following categories: Difficult to Apply/Subjectivity; Types/Levels; Approach to Evaluating; Knowledge of the Staff; and Effort/Value.

Sample Comment, Apply/Subjectivity (19 occurrences noted):
My assessments were somewhat subjective. I'd like to have some sessions to compare notes with peers on how to apply the scale to practice questions to get some common understanding of how to use the scale.

Sample Comment, Types/Levels (16 occurrences noted):
The criterion for each level should have had more concrete benchmarks.

Sample Comment, Approach to Evaluating (9 occurrences noted):
It assumes that a question has an inherent difficulty factor. There is no taking into account the experience or inexperience of the librarian.

Sample Comment, Knowledge of the Staff (6 occurrences noted):
Being uncertain about how effective my rating was when dealing with questions far outside the realm of my normal subject areas—patent questions, etc. would be more complex for me but a piece of cake for our patents librarian. I wasn't sure how to "figure in" that factor.

Sample Comment, Effort/Value (4 occurrences noted):
I also didn't feel like it was clear how to assign a number on the scale when more time than expertise was involved with a reference interaction. The comments around difficulty be-

tween determining Scale levels reflects the outcome of an earlier question, which asked participants to indicate which, if any, categories they had trouble deciding between.

A follow-up question encouraged the participants to suggest alterations to the Scale for future modification. These modifications were put into the following categories: Delivery Method/READ Scale Appearance; Time Element; Skill Level Element; Clarity of Categories; Discussion Component; and Comments/Observations:

Delivery method/READ Scale Appearance (9 occurrences):
Automate it! It would be great to have on the computer.

Time Element (5 occurrences):
Additionally, the numbers in the scale (1–6) may have more meaning and value if time were a factor—I've had "3" interactions that can last anywhere from 5 minutes to 20 minutes, but they are all simply marked "3."

Clarity of Categories/more descriptive/fewer categories (4 occurrences):
The degrees of gradation of reference questions were important, but not very clear. I wish they had been more concrete… like a checklist for each category or more defined descriptions for each category. A revision will reduce the variable/error margin between scoring librarians.… The criterion for each level should have had more concrete benchmarks.

Skill Level Element/experience of reference staff (4 occurrences):
Though it added some context to reference statistics, it could stand a little more context. What may be a 3 or 4 level for someone with little or no experience (a graduate student assistant, for example) may be a 2 or 3 for someone with a great

deal more experience. The scale may have more use if each person at the desk kept [his or her] own statistics, so that experience could be factored in.

Discussion Component/requirement (2 occurrences):
Reference staff should talk openly and often about how to apply question scale levels to make sure we are all on the same page. The descriptions are helpful—but everyone reads things differently. There are gray areas. There are things we all do differently—so I think open discussion would be helpful.

General Comments/Observations (2 occurrences):
The simple nature of the READ scale works to do two contrary things: point out the variability of the work that we do, while showing how limited we are in tracing the ways in which we make knowledge available to each and every patron on an individual level. Statistics, by nature, are too broad and contain not quite enough depth at the same time.

Finally, the study group was asked if their approach to reference changed in any noticeable way during the period they applied the READ Scale to measure their reference work.

The number of the overall participant study group that changed their approach to reference was low, only 10 out of a total 98 responses, but these responses are worth including here, as it provides a snapshot of the online survey participants' range of experience. A small percentage of the participant group indicated difficulty with incorporating the READ Scale into existing reference procedures, while a high level of study respondents experienced more satisfaction, increased awareness, and an appreciation of the effort, knowledge, and skills involved with reference work by applying the

scale to aid in measuring their reference work effort.

I experienced an increased aware-ness of differing levels of reference work.

Frankly, it complicates the process. Trying to delineate between a 1 or a 2, a 3 or a 4, etc., is tedious.

I was more likely to think about the level of service being provided.

I gave more [conscious] thought to the processes or steps involved in order to rate each interaction.

I was more aware of the level of effort that could be applied to ques-tions vs. what I actually did.

It made me keep statistics regularly.

More aware of time spent on transaction(s).

I think I worked a bit harder to make sure that I recorded everything.

I had to think about the level of effort.

I was more self-conscious of the level of help I was providing, with the net result that interactions improved. My level of empathy and understanding (dare I say "patience") improved along with it.

Using READ Scale Statistics: Practical Approaches

The READ Scale was developed as a tool for capturing vital supplemental quali-tative statistics when reference librar-ians assist users with their inquiries or research-related activities by placing an emphasis on recording the skills, knowl-edge, techniques, and tools used by the librarian during a reference transaction.

The researchers propose that there are a number of practical approaches to us-ing the statistical data derived from the READ Scale for both strategic planning and the assessment of reference services. Individual institutions can use READ Scale statistics for staffing; training and continuing education; renewed personal and professional interest; outreach; and reports to administration.

Staffing
Comments from the study:

We've always known empirically that a large percentage of our refer-ence transactions were quick and easy. This study provided concrete evidence of this, with possible staff-ing implications.

It shows a much clearer picture of what we are actually doing with ref-erence. It is possible to see where the true "busiest times" are in the day.

By using the READ Scale, it is possible for libraries to alter staffing patterns to best serve the users and librarians. One institution involved in the study decided to "let go" of requiring full-time profes-sional librarians to staff their reference desk in the mornings and on Saturdays after viewing the number of level 1 and 2 questions they received at those days and times. This empowers student workers and part-time staff, who took over some of the duties, and frees the professional librarians to concentrate on liaison and collection development duties. Another li-brary in the study is using the data to pro-pose reducing faculty librarian scheduled hours in the evening by ending them at 9 pm instead of 11 p.m., having noted that, after 9 p.m., transactions not only become infrequent but are rarely ranked above category 2 on the READ Scale. Prior to using the Scale, the evidence for changing schedules could only be described as anec-dotal. By the same token, the opposite can be noted—high traffic times or notations

of higher categories of the READ Scale can be used to supplement and strengthen the value of reference desk staffing.

Training/Continuing Education

Comments from the study:

I felt it was very useful because it challenged me to come up higher in those areas where I need improvement in certain concentrations like history, which is not my specialty. I need to learn so much more.

Not directly related to the READ Scale itself, but based on the compilation of answers for the sample questions, we realized that not all our librarians were approaching questions in the same way. The ratings could vary from 2 to 6 for the same question. Based on that, we have decided to bolster our staff development and training program and improve our mentoring of new librarians.

The READ Scale can be used as a training tool for librarians at all levels. The second observation above is a great example of how using the READ Scale can assist in the training and mentoring of reference staff. Another service point also reported the same experience—they will also increase training. The researchers suggest that this training can be done throughout the semester or year using the READ Scale. If, at the beginning of the training period, scale effort levels recorded and the answers provided are not in line with each other, a training regimen with outcomes can be developed, and a similar series of questions can be tested at a later date to ensure that the staff is developing the necessary reference skills and knowledge.

As another study participant observed, using the READ Scale encourages continuous learning. The researchers suggest that reference staff could make the most of this opportunity by writing down any

questions that elicit an assignment of a category of 4 or higher on the READ Scale at the reference desk and then sharing these questions and how they were answered with their colleagues, providing the opportunity to discuss strategies for assisting users, and learning from colleagues who have in-depth subject knowledge in that particular area. This could also be a great way of reconnecting with others, for the love of the job. Gerlich's case study reveals that the number-one reason reference librarians chose their profession was to help people with research; the second reason was the aspect of "the detective work."[13]

Renewed Personal and Professional Interest

In Gerlich's case study, reference staff and administrators acknowledged the primary function of their profession as that of providing reference service;[14] likewise, they recognized that current data-gathering methodologies were not sufficient in recording the importance of this work or effort.[15] The READ Scale provides a way of revealing and counting important supplemental data that have been hidden in the customary tick marks used to record reference statistics.

Comments from the study:

Using the READ Scale added to my sense of accomplishment!

The thought required to rank questions according to the READ scale made me think a little at the completion of the reference interaction—and thus to become more self-aware.

It gives ME a tangible scale on which to rate my efforts, ultimately spurring me to strive for better service.

By using the READ Scale, reference staff can rate their effort and receive acknowledgement for their effort, knowledge, and skills as appropriate. The level of skill is especially important to note in a situation where subject or liaison

practices are the norm and librarians are sought out for their expertise and consultation services. In-depth specialized transactions often happen away from the traditional service desk, and credit for expertise is often not recorded or acknowledged.

Outreach

Using the READ Scale can help develop outreach activities for librarians. In an instance where a liaison program is strong, but there is little visible research or library activity and low or no in-office consultations, this may be a sign that outreach efforts should be increased. This would be especially pertinent in an environment with research-intensive programs, where reference staff could expect to assist faculty or upper-class students who would be expected to have intensive assignments, to conduct research or need primary research materials. An active campaign or meeting with the department could elicit an increase in the types of interactions that would be assigned level 4 or higher on the READ Scale.

The same can be said for reference desk statistics in general. If libraries are only experiencing inquiries that require efforts at the 1 or 2 READ Scale categories, then how are students and faculty getting their information? Do they know what services and resources you have? Are there new ways to market services, facilities, or research assistance? Are there times of the year when higher READ Scale categories are showing up in the statistics, and, if so, can those patterns be predicted and assignments be noted to facilitate new research guides, make connections to teaching faculty, or influence new designs or products?

Reporting/Statistics

The READ Scale is intended to record supplemental statistics alongside the traditional quantitative data gathered that could be used by administrators to report the knowledge and skills used in reference services.

Comments from the study:

I liked that it attempts to record the intensity of the reference transaction. In my view that was a sorely missing piece of information when recording in the traditional fashion.

The READ Scale is an assessment tool that does a better job of reflecting how reference librarians spend their time. It gives more value than tick marks on a page. It's a tool we can use with administrators to show what we really do.

Just as READ Scale statistics can help determine staffing strategies, the qualitative nature of the instrument can help with the creation of narrative text more descriptive in nature when developing reports to stakeholders, especially where an administrator needs to explain roles or job functions. This could be particularly meaningful in cases when off-desk statistics are recorded and reference librarians track communications, research assistance, and appointments with their constituents via e-mail. More time and effort are required for those activities but are rarely recorded.

The READ Scale could also be useful in estimating average time spent helping patrons. In the testing phase of the study, participants were asked to record the amount of time it took to complete a transaction. These data enabled the researchers to make rough estimations on the average length of time per transaction for each scale category.

Table 12 illustrates the total number of transactions per category and the estimated number of hours or days needed to complete a transaction, based on the pre-study question calibration data where participants recorded time expended to answer the test questions. These figures can only be used as an illustration of what keeping track of time can be used for, as these data were gathered from the test period and therefore do not take into

account "real time"—that is, time spent talking with a patron, the time involved in conducting teachable moments, the learning skill of the recipient, and so forth—times averaged for these transactions and efforts did not involve a "live" patron. A real transaction, with an interview and resulting conversation, dependent on the needs or communication skills of the user, in all likelihood would have taken longer. Adding a measure of the time expended to handle a transaction was also suggested by some participants in the modification section of the online survey. If a library were to keep track of the time expended for each transaction within a semester, then accurate data could be applied. This is especially useful in terms of real-time electronic services, such as Chat, where the back-and-forth communication takes on a different dynamic than an in-person communication:

> At times, certain aspects of the scale indicating difficulty level seemed to conflict, particularly on [C]hat. For example, there were times when an answer was relatively easy—I knew it based on my knowledge—but because I was working via [C]hat, it required quite a bit of time to guide a user through the information session when I think less time might have been required for an in-person transaction.

Conclusion

Reference staffs appear ready to try new methods for recording reference statistics that include qualifying their effort, knowledge, and skills. By continuing to gather data from institutions that try the READ Scale for reference services, the researchers can begin to amass a large body of statistics to normalize the Scale even more, with an aim to create a dialogue among professionals.

Future Directions for Research

The authors are invested in continuous improvement of the READ Scale and wish to thank our study participants as well as other libraries that have adopted the scale for use at their institutions and for continuing to share their data with us. We have benefitted from users who suggested modifications as well as from having the privilege of being engaged in constructive and fruitful discussions toward progress in improving the measurement of reference work. In our quest to share the READ Scale and to investigate its viability, several aspects of the scale have emerged as elements worth considering for future research.

The most popular inquiry to arise when discussing the READ Scale is the issue of the timing of each category (for instance: on average, how long does a scale 3 question take to answer?) The researchers have considered the element of "time" as a measurement category and

TABLE 12
Final Total Number of Transactions Recorded, Using the READ Scale and Average Time Spent Total (Based on Common Pre-Study Q & A)

Average Time (in minutes)	1	5	7	15	90	90		
READ SCALE	1	2	3	4	5	6		
Service Points	9,497	5,622	3,085	926	303	68	Total	19,501
Off-Desk	658	635	565	295	117	53	Total	2,232
Totals	10,155	6,257	3,650	1,221	420	121	Total	21,824
Hours	169	521	426	305	630	181	Total	2,232
Days (24 hrs)	7	22	18	12	26	8	Total	93
Days (8-hr day)	21	65	53	38	78	22	Total	277

encourage adoptive institutions to build a timing element into their preparation and calibration training tools for librarians gearing up for using the READ Scale in their reference work. We have observed two dominant schools of thought on the proposal of using timing as a continuous measurement. One school favors the timing of each transaction to later be used as a performance measurement tool, a training tool for the calibration of level of effort when applying the READ Scale rankings, and for reporting workload effort to administrators. The opposing school of thought does not favor the use of "timing" as a measure of reference effort as it can vary widely due to the knowledge, experience, and personality of the librarian handling the transaction. The issue of the value of timing reference transactions bears future investigation.

Survey feedback teased out the question of how to take into account an individual librarian's level of reference experience and expertise brought to the reference transaction and how to score for varying levels when "rating" a transaction using the READ Scale. The question of "level of experience and the rating of reference transactions" is an area that would benefit from future research. How does one build in expertise and knowledge that is unique to the librarian or staff member, their familiarity with the resources and policies of their institution, when using the READ Scale? More work remains to address this aspect of the application of the scale.

Determining the effectiveness of the READ Scale for recording reference statistics and applying assessment practices requires continued, long-term data collection from a variety of institutions. The researchers welcome any and all interested libraries to try the READ Scale and contribute to its ongoing development as a supplemental tool for qualifying reference statistics by participating in the ongoing research collaborative and sharing experiences with colleagues. For more information, go to http://www.dom/edu/library/READ/index.html.

Notes

1. Eric Novotny, *Reference Service Statistics & Assessment (ARL Spec Kit #268)* (Washington, D.C.: Association of Research Libraries, 2002).

2. Bella Karr Gerlich and G. Lynn Berard, "Introducing the READ Scale: Qualitative Statistics for Academic Reference Services," *Georgia Library Quarterly* 43 (Winter 2007): 7–13.

3. Brian Quinn, "Beyond Efficacy: The Exemplar Librarian As a New Approach to Reference Evaluation," *Illinois Libraries* 76 (Summer 1994): 163–73.

4. John C. Stalker and Marjorie E. Murfin, "Quality Reference Service: A Preliminary Case Study," *The Journal of Academic Librarianship* 22 (Nov. 1996): 423–29.

5. Jennifer Mendelsohn, "Perspectives on Quality of Reference Service in an Academic Library: A Qualitative Study," *RQ* 36 (Summer 1997): 544.

6. Bella Karr Gerlich, *Work in Motion/Assessment at Rest: An Attitudinal Study of Academic Reference Librarians, A Case Study at Mid-Size University (MSU A)* (2006).

7. Deborah B. Henry and Tina M. Neville, "Testing Classification Systems for Reference Questions," *Reference & User Services Quarterly* 47 (Summer 2008): 364–73.

8. Russell F. Dennison, "Usage-Based Staffing of the Reference Desk: A Statistical Approach," *Reference & User Services Quarterly* 39 (Winter 1999): 158–65.

9. Debra G. Warner, "A New Classification for Reference Statistics," *Reference &. User Services Quarterly* 41 (Fall 2001): 51–55.

10. Novotny, *Reference Service Statistics & Assessment (#268)*.

11. Sarla R. Murgai, "Reference Use Statistics: Statistical Sampling Method Works (University of Tennessee at Chattanooga)," *The Southeastern Librarian* 54 (Spring 2006): 45–57.

12. Novotny, *Reference Service Statistics & Assessment (#268)*, 10.

13. Gerlich, *Work in Motion/Assessment at Rest*, 122.

14. Ibid.

15. Ibid.

"A Declaration of Interdependence"
Philadelphia, March 30 – April 2, 2011

acrl 2011

peer revered

CALL FOR PARTICIPATION:
Great ideas. Multiplied.

Explore new ideas, engage in active peer to peer learning and share your latest research, cutting edge practices or innovative developments with your colleagues.

Session Formats
Deadline May 10, 2010
- Contributed Papers
- Panel Sessions
- Preconferences
- Workshops

Deadline November 1, 2010
- Cyber Zed Shed Presentations
- Poster Sessions
- Roundtable Discussions
- Virtual Conference Webcasts

Submit your proposal online today!

www.acrl.org/acrl/nationalconference

ACRL 2011 registration opens May 2010.

ACRL 2011 Conference