

Data Speaks Volumes:

Evidence-Based Delivery of Library Services in a User-Centered Library

Abstract

With the ever-increasing focus on the needs and experiences of our users, librarians are compelled to reposition themselves as providers of services and support. Libraries are now more than the sum of their collections, and are becoming hubs for creation, instruction, and research. With this increasingly service-oriented approach to the profession, librarians must view their work and services through lenses of critical inquiry, assessment, and thoughtful design. Above all libraries are increasingly expected to plan, deliver, and assess their services based on evidence and data. This evidence-based approach to librarianship demands greater attention in professional literature and best practices.

This paper consists of a review and discussion of evidence-based practices in librarianship, as well as a case study of one academic library's project to assess reference programming using systematic data collection and analysis. Academic librarians in a small Sino-Foreign university situated in cosmopolitan China explored demand for and use of reference services among their users, a diverse and multinational population of undergraduate students and faculty. Data collected included traffic patterns and circulation activity as well as a system to categorize reference transactions by genre. The resulting data-driven system was used to guide the library's scheduling and staffing of in-person and

virtual reference services. In addition, it was a key tool for training library staff and interns for reference work. As one of the library's first attempts at planning, delivering, and assessing reference services to a small but highly engaged population of academic users, this approach exemplifies the successful use of evidence-based planning in a user- and service-oriented culture.

This research is part of an emerging trend in librarianship to rely on evidence and data, rather than opinion and anecdotes, to guide the planning, delivery, and assessment of library services that are truly user-oriented. Many of the themes this conference emphasizes are expressed with evidence-based or service-centered language. This presentation of how to use evidence to guide library programs will connect well with the "Using Evidence" theme as well as "Demonstrating Value" and "Measuring Quality and Success". It contributes to the field by situating a real-life case study within the context of other scholarly works on this topic, which is an area of study that requires more research.

Introduction

Modern academic libraries are repositioning themselves as hubs for creation, instruction, and research, and are increasingly focused on the needs and experiences of users. Librarians are compelled to leverage their expertise in research, instruction, and technology to support users' research needs. As service providers, librarians must view their work through lenses of critical inquiry, assessment, and thoughtful design. Above all, they are expected to plan, deliver, and assess their services based on evidence and data.

Evidence-based assessment strategies inherently support the service-oriented nature of librarianship. Qualitative and quantitative data allow librarians to evaluate the benefits of services provided within the context of user needs, behaviours, and expectations (Bowron & Weber, 2017).

Librarians have kept reference statistics, particularly classification systems for reference interactions themselves, for decades. In her literature review on classification systems for reference statistics, Debra Warner wrote that libraries were interested in assessing the success of reference transactions and correlating reference and instructional work; but when the Association of Research Libraries (ARL) first embarked on its reference statistics project, 23 percent could not supply such statistics (Warner, 2001).

The goal of this paper is to review existing literature on evaluating reference services, with particular focus on the qualitative and quantitative statistical evaluation and coding of reference services and interactions. This theoretical framework is the context for the subsequent case study of one

academic library's experiment using qualitative statistics to understand patron needs and behaviours in order to effectively staff reference service points and tailor library services to meet patrons' demonstrated expectations.

Literature Review

There is considerable research in library literature on strategies for the collection of qualitative and quantitative reference transaction data, and a librarian's choice to use or adapt any of these models is determined by the questions that a library team seeks to answer. Statistics might be used to explore a range of issues such as:

- understanding common challenges or sources of frustration, or "pain points," for patrons navigating library facilities and web services;
- measuring the frequency with which library services and resources are used;
- determining popular service points;
- identifying peak times and days for patron traffic;
- describing language used or expected by patrons in different reference environments;
- gauging the complexity or difficulty of questions, including the expertise required or time spent answering them;
- identifying a collection's strengths or weaknesses;
- classifying types questions received; or
- measuring patron satisfaction, perceptions, or demographics.

There are many methods of collecting this information, including surveys, focus groups, ethnographic studies, transcript analysis, and manual coding systems (Luo, 2008).

Historical Approaches to Collecting Reference Statistics

Writing on the history of the measurement of library services including reference work, Lawrence Thomas lamented the crudity of evaluation methods employed by libraries (Thompson, 1951). He observed that early statistics and reports kept since the nineteenth century overwhelmingly dealt with the content of library holdings. One such example is an early report by Edward Edwards which tabulated the population served by major European and North American public libraries and the growth of their collections over 10 years (Edwards, 1848). As library collections evolved to include digital materials and library services were adapted in response, the focus of library statistics also changed.

In 1998, Carol Tenopir wrote in *Library Journal* that “traditional statistics that count number of interactions with patrons may not reflect...increased activity” and that they failed to explain “what resources are most helpful or what formats require increased time with users” (Tenopir, 1998). To understand the value of these resources and to improve library services, especially while some institutions were grappling with budget restrictions and reduced staffing, librarians turned to more descriptive data (Rawson, Davis, Harding, & Miller, 2013; Sullivan, Schoppmann, & Redman, 1994). Reference data can also help librarians communicate more effectively with users and staff service points appropriately (Bishop & Bartlett, 2013; Hughes, 2014; Scales, Turner-Rahman, &

Hao, 2015) or address needs for new or enhanced support with library services (Arnason & Reimer, 2012). Bishop and Bartlett's research allowed them to use reference staff more efficiently by deploying subject librarians to service points that received the majority of research-intensive questions.

Analyzing and coding reference transactions can also empower librarians to evaluate the value and use of library services. Maximek et al analyzed their transactions to understand who was using their reference services, and why, to determine if their virtual reference work was adequate and matching standards for in-person reference (Maximiek, Rushton, & Brown, 2010). Librarians at Michigan State University Libraries used transaction analysis and Web analytics to investigate user behaviour and evaluate the effectiveness of their service points for distance library users in order to reflect the value of their services (Tobias & Blair, 2015). Web analytics gave librarians insight into how users were interacting with the library website. This was supplemented by a transaction analysis project, whereby librarians analyzed virtual reference chat logs to code questions in order to understand user's pain points and expressions of frustration. Librarians at the University of Tennessee Health Sciences Library conducted a similar content analysis project as part of their website redesign initiative (Fan & Welch, 2016). Analyzing contents of emails, text messages, and live chats, librarians categorized questions to improve the structure of the library website and to understand the language patrons used so that the content of the website was logical and understandable to them.

Lorna Rourke and Pascal Lupien also conducted a content analysis project to collect data on the language patterns of patrons through two service points (Rourke & Lupien, 2010). They found that users used these platforms differently, bringing more research-intensive questions and formal language to the Docutek Virtual Reference¹ service, which and used more informal language on chat services where they tended to pose directional and policy-based questions.

Categorization Schemes for Qualitative Reference Statistics

The categorization of reference questions is a common strategy for library statistics. Neville and Henry pointed out that “reference statistics have been troublesome to collect because of the qualitative nature of questions and the difficulties in assigning a wide variety of these questions into a minimal number of strict categories. The classification of questions must be distinguishable and consistent” (Neville & Henry, 2009, p. 372). Much research has been dedicated to the development of these categories (Baker, 1991; Calabretta & Ross, 1984; Matthews, 2018; Rothstein, 1990; Witucke & Schumaker, 1991), but systems developed by three scholars are especially notable: those by William Katz, Debra Warner, and Bella Karr Gerlich and G. Lynn Berard.

Katz presented his four categories for reference questions in *Introduction to Reference Work*, first published in 1969: directional, ready reference, specific

¹ Here Rourke and Lupien describe Docutek Virtual Reference, a SirsiDynix Company as it was used for “virtual reference” services, which they distinguished from “instant messaging” or “chat reference” services.

search questions, and research questions (Katz, 2002). Subsequent researchers noted that these categories do not capture the complexities of reference questions and that they were inadequately defined. Warner presented a new classification system in 2001 to account this ambiguity (Warner, 2001). Warner's four categories focused on how resources and skills were used during a reference transaction: non resource-based, skill-based (also referred to as 'how-to' questions), strategy-based, and consultation. When Neville and Henry compared the Warner model to traditional classification systems, it was found to be more consistent and easier to use with less confusion about the definition of categories themselves (Neville & Henry, 2009). At the time of their writing, they acknowledged that other libraries at the time were experimenting with a third model developed by Gerlich and Berard: the Reference Effort Assessment Data (READ) Scale.

Development and Applications of the READ Scale

The READ Scale was introduced in 2007 as a six-point measure for classifying reference transactions based on the effort and time required to answer a given question. The authors observed that library literature indicated a decline in transaction numbers and sought to resolve the discrepancy between the number of reference questions handled and the level of expertise needed to answer them. They remarked "this method is especially appealing in a profession where the industry standard for recording statistical data is a hash mark that records and rewards quantity as opposed to quality" (Gerlich & Berard, 2007). Each level of the READ Scale is numerically represented along a scale of 1-6,

with 1 representing reference questions requiring the least amount of time and effort, and demanding no specialized expertise. Level 6 questions require the most effort and time, and include inquiries that cannot be answered on the spot. Introducing the READ Scale, Gerlich and Berard described the process of testing its ease of use and clarity among librarians at Carnegie Mellon University. The authors acknowledged that librarians were generally in agreement when rating questions at levels 1, 2, and 6, but were more likely to differ when ranking levels 3, 4, and 5.

The READ Scale's recognition of library workers' time, effort, and expertise has great appeal, and has been used or adapted by many library systems. Gerlich and Whatley have continued to research the READ Scale and reflected on early use of the READ Scale and its value in determining staffing schedules (Gerlich & Whatley, 2009). Librarians at Austin Peay State University Library successfully incorporated the READ Scale into their assessment methods (Bowron & Weber, 2017), and librarians at the McConnell Library at Radford University used a modification of the READ Scale to train student workers with the goal of encouraging responsible referral practices (Vassady, Archer, & Ackermann, 2015). The challenge of training library workers in proper referral practices is familiar to many librarians. Warner quotes Elizabeth Malia's remarks that their "biggest problem lies in the area of making referrals to the librarians" (Warner, 2001).²

² Originally in Elizabeth Malia, "Triage and Tiers: Paraprofessionals in Reference Service at Eastern Washington University Libraries," *Library Mosaics* (1996).

The READ Scale is available for any institution to use at no cost, and its creators will support interested librarians as they adopt or adapt the READ Scale. As a flexible model, the READ Scale can be integrated with a library's existing system for tracking statistics. The Library at Colorado State University-Pueblo combined their use of the READ Scale with their own web-based program for maintaining statistics, called DART (Hudock & Sullivan, 2011). DART reflected five Student Learning Outcomes (SLOs) (Key Services, Research Tools, Search Strategies, Evaluation, and Citation/Documentation), which were mapped to question categories. For example the "Key Services" SLO would break down to categories such as circulation, book renewal, interlibrary loan, and library hours. Within DART, librarians could augment the five SLOs and their question categories with the READ Scale, allowing librarians to cross-reference a learning outcome with a question category and READ Scale ranking.

Gerlich and Berard have continued to test and assess the viability of the READ Scale, and in one assessment it was found that the majority of library workers using the READ Scale would recommend its use (Gerlich & Berard, 2010).

Case Study

This culture of collecting qualitative and quantitative data has empowered libraries to provide user-centered services and develop evidence-based strategies for staffing, training, and assessment of services. Existing models, such as the READ Scale and those developed by Katz, Warner, and others have been used to great effect at many institutions. Others may find it useful to adapt

these models or develop an entirely new model to address specific research and analytical questions about staffing, resource use, or patron behaviours. This paper focuses on the latter scenario, where an academic library developed an in-house system for classifying and tracking reference transactions to facilitate training, staffing, and the appropriate provision of services.

Many aspects of this institution are unique, and librarians must rise to meet the needs and expectations of their user population. Located in the cosmopolitan heart of a major city in China, this library serves a diverse, multinational and multilingual community of over 1200 undergraduates, faculty, and staff. Instruction at this university is conducted in English, and the majority of the library's collections are in English, with a significant number of Chinese-language resources included. As a joint Sino-Foreign university the library follows policies and service strategies similar to those of North American institutions, which may be new or unfamiliar to non-Western students. Reflecting the diversity of the university and its student body, library staff also represent a variety of nationalities, languages, and expertise. The majority of library workers in this institution are native speakers of Mandarin Chinese, speaking English as a second language.

Further complicating the nature of the library's services, it is part of a globally networked university system with three degree-granting campuses across the globe and a network of 11 global academic centers on five continents. Library workers at each of the three main campuses provide in-person reference services and research support to their local users; these services are augmented

by the provision of virtual reference through email (using Qualtrics and Questionpoint forms) and live chat services hosted by Libraryh3lp. Through this consolidated effort by librarians, virtual reference services are available almost 24/7 to affiliated students, faculty, and staff across the globe.

Since the founding of the university's campus in China in 2012, the library has sought to provide high-quality and readily accessible reference services virtually and in-person. The university, including the library's services and collections, was initially located in a space shared with another local university until it relocated to its own permanent facility in 2014. Through this period, and during the first four years of the library's existence, the availability and visibility of these services have increased significantly.

The library has always provided both in-person and virtual reference support. Librarians staff in-person reference shifts with virtual reference shifts staffed by librarians as well as library assistants and qualified interns. In-person reference was initially provided at the circulation desk, but hours were not consistently scheduled and staffed. As a primary service point, the circulation desk was busy with student workers, librarians, and library assistants. Patrons often expressed confusion about whether librarians were there for research support or to provide circulation services. In response, a dedicated Reference Desk was added to the circulation desk in April 2016. At this same time, librarians began to implement strategies to increase the efficiency of their services and staffing and to better understand the needs and expectations of their users.

Early Approaches to Statistics

Librarians were motivated to develop a method for tracking statistics in order to answer specific questions about user behavior and staffing needs. Later refinements to this project were similarly designed to collect data purposefully, rather than tracking statistics because they might be interesting or simply for the sake of doing so.

Changes in staffing through departures and new hires, a desire to increase the availability and efficiency of library services, and a request from the library's leadership for the library staff in China to cover more virtual reference hours confronted librarians with a need for insightful reference statistics. A collaborative Google Spreadsheet was designed in April 2016 to capture this data. In addition to tracking the number of reference questions handled through Libraryh3lp and in person, this document allowed library workers to maintain a shared record of one-on-one research consultations held in-person and through e-mail, and the number of students reached through pre-scheduled and requested workshops.

In the first iteration of this statistics project, library workers manually tracked reference statistics and circulation desk statistics separately, though both were quantitative measures of the number of questions received in each shift. When tracking interactions at the circulation desk, library workers would indicate both reference questions and circulation statistics manually with a check mark. Similarly, the reference statistics form only indicated the number of reference questions handed, with a column included for additional notes and observations.

The purpose of this notes field was for library workers to describe particularly difficult or unusual questions; it quickly became apparent that there was no clear definition of what constituted a “difficult or unusual question”. For example, a question about locating toilets or understanding the library access policies from a user in one of the university’s other global centers would be very difficult for any library worker unfamiliar with that building. As a result, library staff often felt compelled to detail each question received regardless of its difficulty, making the statistics tracking form cumbersome to use and producing little valuable data. Another challenge that library workers encountered was training library assistants, interns, and student workers to know when and how to refer research questions to librarians. Librarians sought to cultivate an environment where referring questions was common practice and encouraged. In this way, library assistants and interns would be empowered to handle questions responsibly and with confidence.

Librarians also hoped to staff their reference shifts more efficiently based on statistical data. Bowron and Weber remark that “when outlining reference desk work schedules, it is useful to identify times of highest patron activity and needs for assistance” (Bowron & Weber, 2017). However, like Bowron and Weber, library workers recognized that there is not necessarily a correlation between the number of questions received and the level of complexity or difficulty of those questions.

Beginning in the Fall 2016 semester, a librarian would simultaneously staff the reference desk and chat reference queues eight hours a week. Before this, a

librarian would only staff the reference desk during her shift while another librarian or library assistant covered the chat reference queues. The decision to consolidate effort was based on quantitative transaction statistics and on personal observations that neither service point was exceedingly busy. Although those statistics do not indicate how many of those transactions were advanced research questions, anecdotal reports from library staff and users indicate that this heightened visibility has encouraged patrons to approach librarians with research questions. Additionally, it was felt that patrons were less likely to approach librarians with circulation questions. In addition to these eight hours of combined reference shifts, the library also provided 7-12 hours of chat-only reference shifts per week.³ Library workers covered these shifts based on individual availability. As a result, library assistants and interns would often receive challenging research questions that should have been referred to librarians, but the lack of clear referral guidelines and practices impeded this. Library leadership recognized that the reference statistics model could be redesigned to allow the reference coordinator to identify patterns in user traffic and staff reference shifts more efficiently based upon statistical evidence.

Developing and Implementing the TARP Model

In December 2016, librarians discovered the READ Scale, which empowered them to reconsider how they viewed and understood reference

³ Although North America observes Daylight Savings Time (DST), China does not. When DST is in effect, librarians in China staffing virtual reference shifts cover 5 fewer hours per week than when DST ends.

transactions handled virtually and in person. The system of categorizing questions based on certain characteristics was appealing and librarians worked together to analyze the READ Scale for its potential applicability to their own institution. As librarians discussed the structure and focus of the READ Scale, some voiced concerns that the six levels seemed overly granular and specific for the nature of the small quantity of questions handled at their own institution. The distinction between the levels of the READ Scale, especially higher levels, also seemed highly subjective. This recalled the existing issue of library workers differing in opinion on the definition of a “difficult” or “unusual” question. The ambiguity inherent in defining complexity was further complicated by differences in language and culture within the library team. Furthermore, the focus of the READ Scale was the perceived effort, difficulty, and time required to answer a reference question; it was the experience of library workers at this institution that the difficulty of a question did not always directly correlate to the amount of time or level of expertise needed to explain or troubleshoot an issue. For its many benefits and qualities, the READ Scale was not an ideal fit for answering the questions librarians were asking, and its subjectivity was prone to conflicting interpretations of difficulty through the lenses of library workers’ cultural or linguistic frameworks.

Instead, the librarians were inspired by the READ Scale to develop their own model for categorizing and tracking reference transactions. Like the systems developed by Warner and Katz this model codes questions by their *type*, not by their perceived difficulty. Librarians named this model TARP, and presented the

model under the headline: “TARP has you covered!” TARP is an acronym that stands for the 4 types of questions tracked:

- **Technical (T)** - helping a patron troubleshoot technical issues with tools & resources such as printers or Wi-Fi.
- **Access (A)** - questions generally considered within the purview of the library’s Access Services team, including accessing spaces, accounts, and document delivery. This term was chosen because these access elements correlated to the specific duties of the access services team, and many library assistants primarily worked as part of the access team. This made it easier to clarify the difference between *Access* questions and *Ready Reference* questions.
- **Ready-Reference (R)** - questions about locating or getting to known research and library resources that do not require the specialized expertise of a librarian. “I know what source I need, but not how to locate it.”
- **Peculiar (P)** - questions about conducting research, citations, and discovering library resources that would usually be referred to a librarian. “I need to discover new sources, identify good databases, or refine my research topic.”

To further explain how TARP might be identified or used, librarians wrote example questions in each category that might be especially or common:

- **T:** “Where are the printers?” “Printers are out of ink.”

- **A:** “What are the library hours” “How do I visit our special collections?”
“Where are the bathrooms?” “Can you help me place a request through Interlibrary Loan?”
- **R:** “I have a citation for an article that I want to read. Can you help me locate it?” “Can you help me navigate call numbers?”
- **P:** “I am working on a research project - what are good databases to use for my topic?” “Can you help me develop a better keyword search?”

Library workers also adapted the statistics methods used at the circulation desk to include the TARP Model. Thus, librarians were able to view in a single document the types of questions received at all access points, and could identify trends and peaks in the number and character of questions.

The TARP Model also proved useful for training new library workers, including professional librarians and library assistants to student workers and interns. A trainer could review past chat transcripts and work with the trainee to discuss how they both would classify a question, and justify that classification. In addition to helping the trainee better understand the logic behind the TARP Model, this allowed the trainer to recognize misleading or unclear parts of the model or otherwise improve it for usefulness and clarity. Librarians developed scripts for common T-type, A-type, and R-type questions for staff to use when answering questions.

Library staff liked the clarity of the TARP Model categories, especially since training materials included explicit examples of P-type questions and built in a system for student workers, paraprofessionals, and interns to know when

and how to refer complicated research questions to librarians. Library workers reported feeling comfortable with this referral system, and believed that they were handling these difficult questions appropriately. This characteristic of the TARP model distinguished it from Katz, Warner, and the READ Scale. Since transitioning from a quantitative statistics tracking approach to a qualitative one in the Spring 2017 semester, the data indicated that the majority of reference questions handled virtually and in person at the circulation and reference desks were T-type and A-type questions.

Analysis of Reference Data

An analysis of the reference transaction data collected for in-person consultations by phone, via chat services, or at the reference and circulation desks illustrated many clear patterns in user behaviour. Library staff answered reference question by phone or in-person at the circulation and reference desks seven days a week, and by chat queues during weekdays in the early afternoon. The least number of questions were asked on Saturdays and Sundays, when classes are not held, and the library has shorter hours. During the week, however, traffic to in-person and virtual service points alike was highest on Mondays, gradually tapering off throughout the week (Figures 1 and 2).

Total Questions Received In-Person, by day

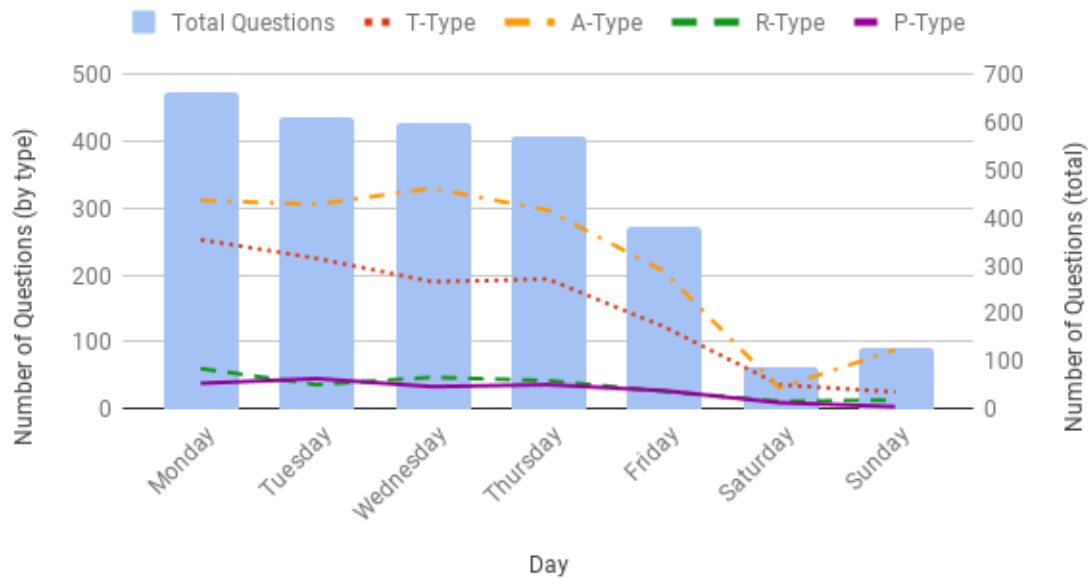


Figure 1. Total number of questions received in-person per day from January 1st 2017 through May 31st 2018.

Number of Questions Received via Chat

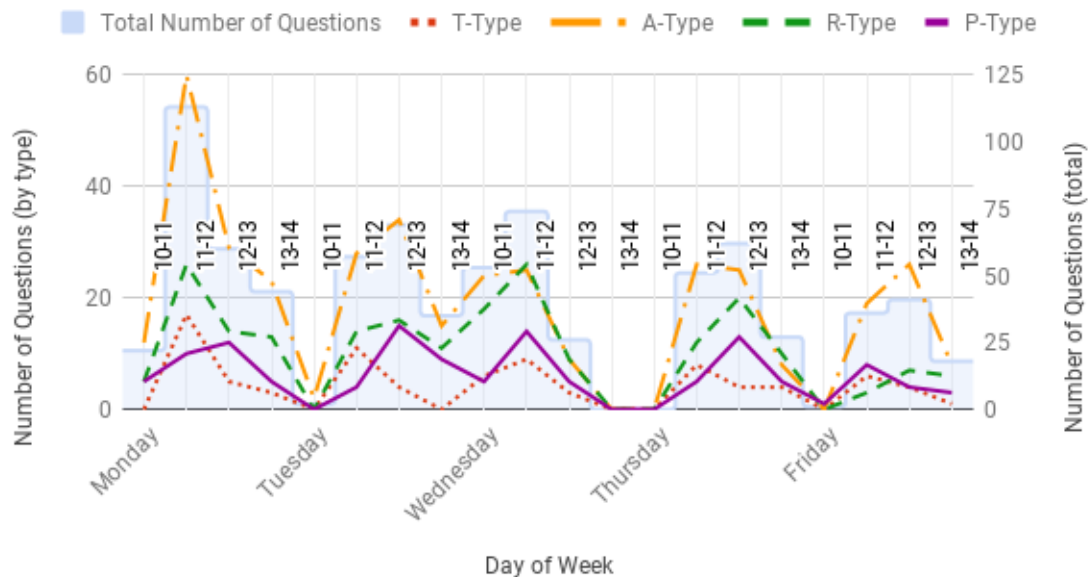


Figure 2. Total number of questions received via chat queues per day from January 1st 2017 through May 31st 2018.

<i>Time</i>	<i>Day</i>	<i>Total Virtual Reference Shifts Staffed with Questions</i>	<i>Total Number of Questions</i>	<i>T-Type</i>	<i>A-Type</i>	<i>R-Type</i>	<i>P-Type</i>
Monday	10-11	7	22	0	12	5	5
	11-12	41	113	17	60	26	10
	12-13	41	60	5	29	14	12
	13-14	41	44	3	23	13	5
<i>Monday Total</i>		131	240	25	124	59	32
Tuesday	10-11	1	2	0	2	0	0
	11-12	24	57	11	28	14	4
	12-13	32	69	4	34	16	15
	13-14	30	35	0	15	11	9
<i>Tuesday Total</i>		87	163	15	79	41	28
Wednesday	10-11	25	53	6	24	18	5
	11-12	39	74	9	25	26	14
	12-13	29	26	3	9	9	5
	13-14	1	0	0	0	0	0
<i>Wednesday Total</i>		96	153	18	58	53	24
Thursday	10-11	1	0	0	0	0	0
	11-12	26	51	8	26	12	5
	12-13	37	62	4	25	20	13
	13-14	32	27	4	8	10	5
<i>Thursday Total</i>		96	140	16	59	42	23
Friday	10-11	2	1	0	0	0	1
	11-12	20	36	6	19	3	8
	12-13	26	41	4	26	7	4
	13-14	17	18	1	8	6	3
<i>Friday Total</i>		66	96	11	53	16	16
Grand Total		476	792	85	373	211	123

Table 1. Total number of questions received via chat queues per day from January 1st 2017 through May 31st 2018.

Virtual Reference questions, which were predominantly initiated by patrons at the university's North American campus, peaked at 12PM. In-person transactions began to increase steadily from 10AM, peaking at 1PM, and decreasingly slowly throughout the before dramatically dropping off at 5PM and again at 6PM (Figure 3).

Total In-Person Questions Received, by hour

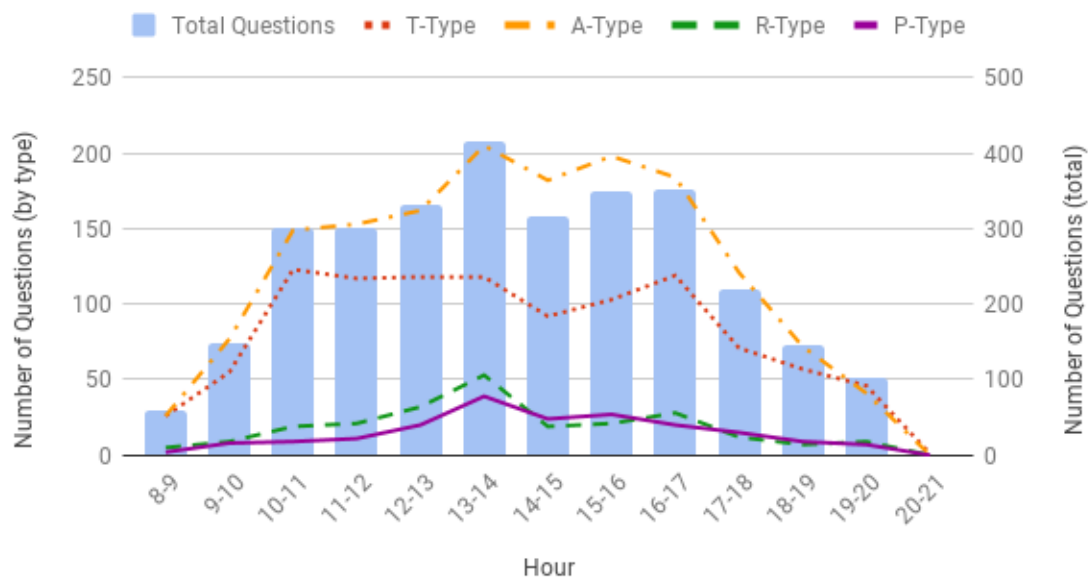


Figure 3. Total number of questions received in-person per hour from January 1st 2017 through May 31st 2018.

Patterns of P-type questions mirrored these overall patterns. P-type questions were most common between 12PM and 6PM, peaking at 1 PM just after the lunch hour concluded, and often became one-on-one research consultations with librarians.

After a year of collecting reference statistics with the TARP Model, the reference coordinator was able to identify patterns in when P-type questions came to certain service points. This allowed her to determine if reference desk hours should be expanded or modified, and to schedule library assistants and interns to virtual reference shifts that typically had fewer P-type questions. Additionally, understanding these trends in user behaviour will guide programming decisions and outreach strategies. Presently, the library team intends to continue using the TARP Model, and is planning to review email transcripts for reference and access-related questions to retroactively add to statistics tracking one-on-one communications, which also uses the TARP Model.

Conclusion

Throughout the process of developing the TARP Model, librarians considered what information they hoped to collect and what questions they needed to answer with this statistical data. Based on these questions and considering potentially complicating factors such as language barriers and varying levels of expertise, the resulting TARP Model suits the library's demands. Librarians found that developing the system to reflect the university's culture made it easier to understand and implement. Useful strategies for developing and adapting models for tracking and categorizing reference statistics include:

- conducting a transcript analysis of chat reference questions to identify common themes or patterns in user language and behavior, or to identify patron pain points.
- evaluating reference questions collectively to develop a common language for categorizing questions;
- defining categories for questions clearly, including example questions or sub-categories;
- researching and adapting existing models, such as the READ Scale or the models proposed by Katz or Warner, to develop a qualitative rubric;
- periodically reviewing the system designed in order to improve unclear elements or update the model to match changes in the library's culture;
- training and refreshing staff as a team in to ensure that staff are defining categories or categorizing questions in harmony.

Approaches to collecting and evaluating statistics and providing evidence-based services are as unique and the institutions that apply them – there may never be a system for collecting reference statistics for all libraries (Neville & Henry, 2009). It is entirely possible that over time an individual library will modify its strategies for collecting statistics, or transition from one model to another based on its changing needs. The next step for library workers using the TARP Model is to review its definitions and categories, retrain staff, measure its effectiveness, and develop or refine user services based on the lessons learned from analyzing the data collected.

References

- Arnason, H., & Reimer, L. (2012). Analyzing Public Library Service Interactions to Improve Public Library Customer Service and Technology Systems. *Evidence Based Library and Information Practice*, 7(1), 22–40.
- Baker, S. L. (1991). *The Measurement and Evaluation of Library Services* (2nd ed.). Arlington, Va.: Information Resources Press.
- Bishop, B. W., & Bartlett, J. A. (2013). Where Do We Go from Here? Informing Academic Library Staffing through Reference Transaction Analysis. *College & Research Libraries*, 74(5), 489–500.
- Bowron, C. R., & Weber, J. E. (2017). Implementing the READ Scale at the Austin Peay State University Library. *The Journal of Academic Librarianship*, 43(6), 518–525. <https://doi.org/10.1016/j.acalib.2017.08.010>
- Calabretta, N., & Ross, R. (1984). Analysis of Reference Transactions Using Packaged Computer Programs. *Medical Reference Services Quarterly*, 3(3), 23–47. https://doi.org/10.1300/J115v03n03_02
- Edwards, E. (1848). A Statistical View of the Principal Public Libraries in Europe and the United States of North America. *Journal of the Statistical Society of London*, 11(3), 250–281. <https://doi.org/10.2307/2337957>
- Fan, S. C., & Welch, J. M. (2016). Content Analysis of Virtual Reference Data: Reshaping Library Website Design. *Medical Reference Services Quarterly*, 35(3), 294–304.
<http://dx.doi.org.proxy.library.nyu.edu/10.1080/02763869.2016.1189783>

- Gerlich, B. K., & Berard, G. L. (2007). Introducing the READ Scale: Qualitative Statistics for Academic Reference Services. *Georgia Library Quarterly*, 43(4), 4.
- Gerlich, B. K., & Berard, G. L. (2010). Testing the Viability of the READ Scale [Reference Effort Assessment Data]©: Qualitative Statistics for Academic Reference Services. *College & Research Libraries*, 71(2), 116–137.
- Gerlich, B. K., & Whatley, E. (2009). Using the READ Scale for Staffing Strategies: The Georgia College and State University Experience. *Library Leadership & Management*, 23(1), 26–30.
- Hudock, S., & Sullivan, D. (2011). Hit the Reference Transaction Mark: Incorporating READ into DART. *University Library, Faculty Publications*.
- Hughes, A. M. (2014). Content Analysis of Reference Transactions Provides Guidance Regarding Staffing of Library Service Points. *Evidence Based Library and Information Practice*, 9(2), 31–33.
- Katz, W. A. (2002). *Introduction to reference work* (8th ed.). Boston: McGraw-Hill.
- Luo, L. (2008). Chat Reference Evaluation: a Framework of Perspectives and Measures. *Reference Services Review*, 36(1), 71–85.
<https://doi.org/10.1108/00907320810852041>
- Matthews, J. R. (2018). *The Evaluation and Measurement of Library Services* (Second edition.). Santa Barbara, CA: Libraries Unlimited.
- Maximiek, S., Rushton, E., & Brown, E. (2010). Coding into the Great Unknown: Analyzing Instant Messaging Session Transcripts to Identify User

- Behaviors and Measure Quality of Service. *College & Research Libraries*, 71(4), 361–373.
- Neville, T. M., & Henry, D. B. (2009). Reference Classification - Is It Time to Make Some Changes? *Reference & User Services Quarterly*, 48(4), 372–383.
- Rawson, J., Davis, M. A., Harding, J., & Miller, C. (2013). Virtual Reference at a Global University: An Analysis of Patron and Question Type. *Journal of Library & Information Services in Distance Learning*, 7(1–2), 93–97.
<https://doi.org/10.1080/1533290X.2012.705624>
- Rothstein, S. (1990). The Measurement and Evaluation of Reference Service. *The Reference Librarian*, 11(25–26), 173–190.
https://doi.org/10.1300/J120v11n25_06
- Rourke, L., & Lupien, P. (2010). Learning from Chatting: How our Virtual Reference Questions are Giving Us Answers. *Evidence Based Library and Information Practice*, 5(2), 63–74.
- Scales, B. J., Turner-Rahman, L., & Hao, F. (2015). A Holistic Look at Reference Statistics: Whither Librarians? *Evidence Based Library and Information Practice*, 10(4), 173–185.
- Sullivan, W., Schoppmann, L. A., & Redman, P. M. (1994). Analysis of the Use of Reference Services in an Academic Health Sciences Library. *Medical Reference Services Quarterly*, 13(1), 35–55.
https://doi.org/10.1300/J115V13N01_03
- Tenopir, C. (1998). Reference Use Statistics. *Library Journal*, 123(8), 32–34.

- Thompson, L. S. (1951). History of the Measurement of Library Service. *The Library Quarterly: Information, Community, Policy*, 21(2), 94–106.
- Tobias, C., & Blair, A. (2015). Listen to What You Cannot Hear, Observe What You Cannot See: An Introduction to Evidence-Based Methods for Evaluating and Enhancing the User Experience in Distance Library Services. *Journal of Library & Information Services In Distance Learning*, 9(1–2), 148–156.
- Vassady, L., Archer, A., & Ackermann, E. (2015). READ-ing Our Way to Success: Using the READ Scale to Successfully Train Reference Student Assistants in the Referral Model. *Journal of Library Administration*, 55(7), 535–548. <https://doi.org/10.1080/01930826.2015.1076309>
- Warner, D. G. (2001). A New Classification for Reference Statistics. *Reference & User Services Quarterly*, 41(1), 51–55.
- Witucke, V., & Schumaker, C. J. (1991). Analyzing Reference Activities: The Affordable Solution. *RQ*, 31(1), 58–69.