

# Lifting the Veil

## *Analyzing Collaborative Virtual Reference Transcripts to Demonstrate Value and Make Recommendations for Practice*

**Robin Brown**

**Robin Brown**, MLS, MA (rbrown@bmcc.cuny.edu) is Assistant Professor, Borough of Manhattan Community College, New York, New York.

*Extended transcript analysis was used to analyze how our chat reference was being used and make recommendations for practice. Because this analysis was longitudinal (over a year, or at least several months) significant patterns were documented. Several themes were noted that emphasized the unique characteristics of the Community College population. The project documented that chat reference patrons are persistent. The questions that came up about assessing the service underlined the commonalities between virtual and face to face reference.*

**B**orough of Manhattan Community College (BMCC) is a large (24K degree-seeking students) urban community college. Students have wide-ranging gifts and abilities and come from 155 countries. The BMCC Library offers 24/7 chat reference using Question-Point from OCLC. Other than scanning the results of a brief satisfaction survey, there has been no attempt to understand how the services are being used. This project used an extended transcript analysis, to describe how the service is being used, to answer questions of value, and to make practice recommendations.

This project was undertaken within the context of an initiative from the

Association of College and Research Libraries (ACRL) to clarify and promote the value of academic libraries. These discussions directly led to creation of the Assessment in Action program.<sup>1</sup> This project was undertaken within the context of Assessment in Action.

---

### LITERATURE REVIEW

Looking at these transcripts beyond the quantitative patterns raised several issues that are confirmed by the literature, including the impact of weak Internet infrastructure on remote services and how that is connected to the fact that this is a commuter school. The challenges offered by college password systems are not uncommon. Do some of the usage patterns reflect our community? Finally, what are some of the possible ways that the relative success of chat reference can be assessed?

Access to adequate broadband Internet within low-income communities is a critical issue.<sup>2</sup> Naylor conducted focus groups to explore why their chat reference was under-used.<sup>3</sup> They observed that lack of “high-speed Internet access” was a very significant issue.<sup>4</sup>

Armann documented the problems that can ensue when the library

password is different from the main college password.<sup>5</sup> This was confirmed by our study.

New York City is a city of immigrants. The New York City Department of Education documents 180 different languages spoken at home by New York City public school students.<sup>6</sup> Conway examined persistence of immigrant students at BMCC: “A third of the total college freshman class spoke English as a second language.”<sup>7</sup> This is an important context for this study.

One of the possible measures of success of a transaction is gratitude. Mon and James examined the percentage of email reference transactions that received a “thank you.” They found that 15.8 percent of the messages received a “thank-you” response. They suggest that it is difficult to conclude that this is absolutely a measure of success.<sup>8</sup>

The challenge with assessing virtual reference is twofold: a level of anonymity (sometimes bridged by identical questions and usernames) and a very short time duration of the interaction. It’s interesting to observe how different research studies have tried to get around these problems. Ward used proxies to ask virtual reference questions, and then used a rubric to grade the “completeness” of the answers.<sup>9</sup> Waugh used interviews to look at whether students have a greater comfort level with formal or informal language.<sup>10</sup> Naylor used focus groups to explore why their service was so lightly used.<sup>11</sup> This suggests that transcript analysis, is only one dimension of examining how this service is used.

There also some new innovations in transcript analysis. Armann-Keown refers to the complexity of the reference questions that were appearing in chat reference.<sup>12</sup> Kemp and Maloney describe the uses of the READ Scale to evaluate the complexity of virtual reference questions.<sup>14</sup>

Passonneau and Coffey (2011) made an important theoretical contribution to “Lifting the Veil.” With their essay they offered a way of tackling a very large amount of data. Grounded theory begins with the data, and develops hypotheses from the data. They also offer a series of questions that can focus the researcher’s attention. The first three are valuable focal points for this study: “1. Contextual: What is happening? 2. Diagnostic research: Why does it exist or happen? 3. Evaluation research: How well does it happen?”<sup>15</sup> This project undertook quantitative descriptions of how things are being used. By looking at what they were asking, and the implied classes, this was a form of contextual and diagnostic research. Evaluative questions are also considered.

---

## METHOD

### The Research Design

QuestionPoint is a collaborative virtual reference service offered as a subscription by OCLC. Borough of Manhattan Community College subscribes to QuestionPoint, as part of a group that includes nine City University of New York

(CUNY) Libraries. This project took advantage of the fact that OCLC makes transcripts available for offline analysis.

These transcripts were received from OCLC, with student contact information stripped out. Grounded theory analysis was very influential in the decision to simply read through the transcripts, and allow the data to dictate the form and the scope of the project. The first read through was the slowest and yielded a rough list of topics and outcomes. Some of the specific inquiries appeared during the first run through. How to quantify the password problem? What about English fluency? Is it possible to measure persistence? The second pass was much faster and tried to code for the class that generated the assignment. A third partial pass looked at issues of correctness and applied an instructional benchmark. A fourth pass (also partial) looked at coding for persistence, to quantify it.

### *Tools for Qualitative Analysis*

This was a “home brew” solution for qualitative and quantitative analysis. MS Excel, with the addition of Power Query and Power Pivot, was an inexpensive solution for this project. This package became less practical as time moved forward. It is only available in Windows, and it is difficult to transfer spreadsheets from one computer to another, and it is impossible to move away from Excel, to Google Sheets. The final two partial scans were repeated using Dedoose ([www.dedoose.com/](http://www.dedoose.com/)), a cloud based platform that made the qualitative coding much easier and much more portable between computers. Dedoose has a graphic interface that makes it possible to tag chunks of text. It is priced moderately on a subscription basis, making it a possible solution for researchers who do not have funding for software purchases. Access to software for doing this kind of work represents a fundamental challenge for faculty doing qualitative research.

---

## RESULTS

In 2013 there were 823 transcripts generated. Both quantitative and qualitative measures were used to try and get a sense of how the service was being used. When was the service used in 2013?

Figure 1 represents the number of transcripts, totals over the year by time of day. What’s interesting about this curve is that it is seen elsewhere on the Internet, as a pattern for shopping, or other forms of customer service.<sup>17</sup> When looking at time of day, it is clear that about 20 percent of our questions are being asked outside of library hours. This quantifies one of the values offered by subscribing to the 24/7 service offered by QuestionPoint.

Figure 2 represents the number of transcripts broken out by day of the week, and then colored for time of day. It’s important to recognize that both of these charts represent total numbers throughout the year. This stresses the point that the service, when compared to the considerable size of

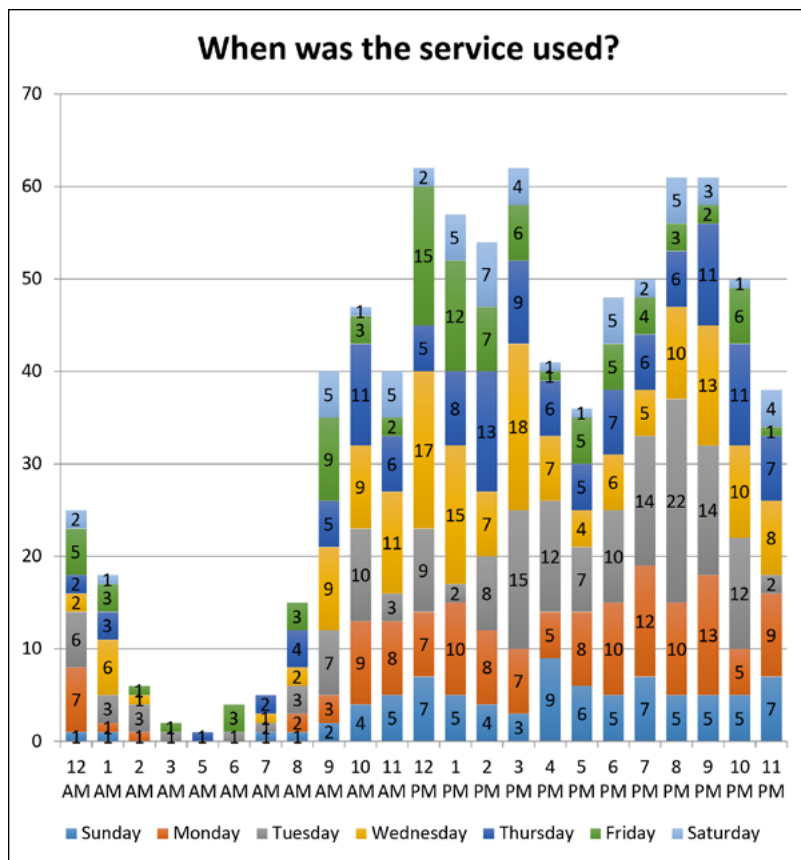


Figure 1. Totals by the hour over a year, color coded by day of the week

Table 1. Who is answering our questions?

QP Backup	246	29.85%
Hunter College, NYC (CUNY)	109	13.23%
CUNY Graduate Center	52	6.31%
Bronx Community College (CUNY)	51	6.19%
Brooklyn College (CUNY)	38	4.61%
Virginia Community College System (VCCS)	26	3.16%
Univ of Hawaii at Manoa	15	1.82%
Lehman College (CUNY)	14	1.70%
Baruch College (CUNY)	14	1.70%
Eastern Michigan University	12	1.46%

the intended audience, is actually very lightly used.

### Who answered our questions?

A glance at table 1 shows several trends. QP Backup represents the overnight coverage, librarians that are directly employed by OCLC (about 30 percent). Hunter, CUNY Grad Center, Brooklyn, Bronx Community College, Lehman, and Baruch College are all within the CUNY group (CUNY

questions appear to CUNY Librarians who are online before they appear to the whole cooperative.) The distinctive contributors to the top ranks are the Virginia Community College System, University of Hawaii at Manoa, and Eastern Michigan University. This documents the important work of a collaborative system.

### What were the questions about?

Trying to summarize the content of these questions was extremely challenging. Two different passes generated an interpretation of topics, and then an attempt to identify which class the assignment was for. Often the patron did not supply much information about what they were looking for, or what assignment was driving their research. Sometimes chat reference librarians, who are not familiar with the campus, don't know what follow up questions to ask.

Some of the most frequent questions included various login problems (see technical problems), content for speeches, and textbook issues. This reflects the fact the Library has a comprehensive textbook collection. All students also take Speech 100, and often are confronted with their first introduction to research when getting ready to make their informative or persuasive speeches. Many Speech professors suggest that statistics are a good way to demonstrate the importance of an issue. This is probably why statistics appeared toward the top of the topic list.

### Disciplinary breakdown

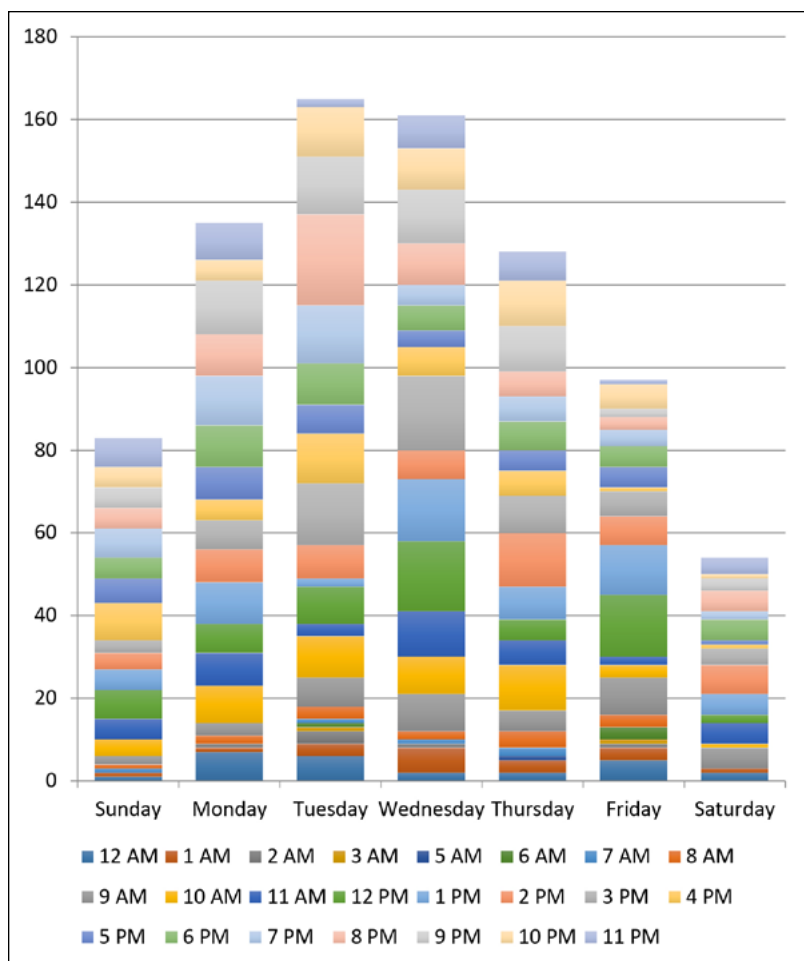
Table 2 represents the most frequent departmental/class specific designations. "Library" questions include questions that were vague enough that it was impossible to discover the class that was generating the question. Since this number represents almost 50 percent of the transcripts, the rest of the numbers must be considered to be "fuzzy" at best.

The English and the Speech classes are the two major places where students are exposed to research projects, and it is not too surprising that they are high in the disciplinary ranking. That year there was a lot of questions about a particular business research project.

Rhetoric was the designation that was used for questions that were about formatting papers and specifics of

Table 2. Specific Disciplines

Library	370
English	108
Speech	102
Rhetoric	29
Business	27
History	15
College	14
Health	12



**Figure 2.** Time of day and day of week

MLA or APA style. This is also not surprising, since the Library continues to be a major source of information on “style issues.”

“College” questions were questions that were referred to another department of the college. Grade interpretation, looking for sample syllabi, and questions about paying for textbooks with financial aid are some examples.

The college does have health courses. Health questions may also represent nursing courses or speeches on health topics. This is an example of why it was so difficult to code for courses or departments.

## OBSERVATIONS

### Persistence

This study found that virtual reference patrons are persistent. In a five-month sample (January–May 2013), 10 percent of the transcripts were judged to be following up on previous chat sessions. This was determined by identical user names, and phrases and questions within a limited

period. Despite the lack of email addresses, persistence jumped off the page. Because of the distributed nature of the cooperative, often the librarian who picked up the question was not aware that this question had already been asked at least once (several students came back multiple times). Sometimes they would get a different, more coherent answer, and sometimes they came back repeatedly and received the same answer—until the explanation sank in. A few clearly came back with connected questions over several days (identifiable by particularly distinctive user names). This is a unique observation, and may be a much larger phenomenon than is possible to document without users being specifically identified and tracked.

### *Signs of Major Technical Problems for the students*

There was seeking help with specific, direct problems with using library services. Often it was about having trouble logging into the databases remotely. At that time, a lot of students had trouble remembering their login password, because it was not a major password at the college that they needed to use on a day-to-day basis. Although it’s hard to completely pin down, 17.4 percent of the questions were judged to have a login component.

Students were having a lot of trouble with their Internet service. There were many complaints about “slow Internet.” If a session crashes, is it the technical skill of the student, or the quality of their Internet service? We attempted to identify the location of IP addresses and got some odd results. There were a lot of local sounding questions with out of town IP addresses. How stable are (or were) the cheapest Internet Service Providers?

### *English fluency*

It was the reaction of this researcher that many of the transcripts reflected English fluency issues. It’s hard to calibrate, because the issue is very subtle, and impossible to measure. Sometimes the chat format may be blamed for a perceived lack of fluency.

### *Success*

Success is a tricky question, particularly when dealing with a moment in time. Wrestling with the content of these transcripts, gratitude was selected as somehow indicative of the “success” of the transaction. About 70 percent of the transcripts showed some level of gratitude. There was no

---

## FEATURE

certainty that this had any real meaning. Some people were overwhelmingly grateful, others seemed to be saying thank you out of politeness. It's impossible to be certain, but it was the perception of this researcher that gratitude is not a reliable measure of success. Without the opportunity for follow up, it's impossible to determine how accurate or successful the answers are.

---

## DISCUSSION

One of the basic ideas behind this study was to demonstrate value and then to underline issues of practice. The numbers demonstrate a service that is being used 24/7 and is available at point of need. The quantitative measures show that the service is being used by students as needed, with a cyclical increase as the semester progresses. The success of the collaborative virtual reference project is evidenced by this data.

*Persistence* is a critical finding of this study. We found that around 10 percent of the sample represented students who were returning to the virtual reference portal, sometimes multiple times. This is a soft number. It could be higher. It was tracked based on noticing unique user names, coupled with identical questions. If students logged in with different or generic user names, they would not have been tracked. This also did not attempt to identify students who were returning throughout the semester for help on different projects. This is likely, but it was impossible to be definite with it. This was due to working with data where the email addresses had been stripped out.

*Technological barriers* jumped out of these transcripts. How stable is their technological infrastructure when doing research from home? This again reaches back to the specific circumstances of this particular population. This is a commuter school, and many of the students are first generation college students. In Fiscal 2013 about 80 percent of students applied for financial aid. Of the students who applied for financial aid, 70 percent were either low income or at poverty level.<sup>17</sup> They are economically challenged. It is possible that their technological infrastructure away from campus is going to be less stable, and less reliable. What is unknown is whether this is still true. There has been less evidence of technological instability in current transcripts.

This study showed that a significant portion of the virtual reference traffic was generated by students having password problems. Since this sample was taken, it's possible that this problem has resolved, since the "library password" is now also the "wireless password" and the "pc login password," which is an improvement from 2013.

*English fluency* is an important issue to discuss. It is difficult to pin down in the anonymous environment of chat reference, but it was definitely present in this dataset. It leads to speculations about the chat reference tool being attractive to students who are concerned about asking for help face to face. Like the sense of technological instability, this is an important window on the particular characteristics

of community college students, and the important role of community colleges in helping immigrants move forward. Chat reference should be seen as an important platform for serving these students.

### *QuestionPoint Discussion*

One of the fundamental challenges of consortium-based chat reference is encountering different styles of reference. Many of the chat reference transactions are resolved with more information and less instruction. It's a truism in academic work that we do not dictate how someone provides reference service, but it's natural to wrestle with whether "I would have done it differently." The answers were not incorrect, usually, just a different approach.

One of the practices on QuestionPoint is that a transcript can be referred back to the student's school, at the librarian's discretion. One of the first takeaways of this project was the realization that many transcripts that would benefit from additional attention are not referred. **Review** of all the transcripts generated by chat reference has become a daily practice in our library.

Each institution that subscribes to QuestionPoint has the opportunity to load a "policy page." The use of the **policy pages** is critical for librarians working within the context of a collaborative service. The correctness of a reference answer is always difficult to judge, but in this study most of the true errors of fact were problems that could have been prevented by reading the policy page.

Do we encourage *extended consultations*? There is an important benchmark within the QuestionPoint community that measures participation by each organization by the number of questions that are answered. There were many transcripts in this dataset that were quite long, as the librarian patiently worked through challenges with an individual student. Collaborative services need to find measures of participation that acknowledge this balance between quantity of questions and time spent with individual students.

---

## CONCLUSION

We were successful in drawing a multidimensional picture of how our virtual reference service is being used. The value of this service has been specifically measured. We noted significant usage outside of library hours. This confirms the basic premise of subscribing to a collaborative 24/7 service. This study absolutely demonstrates the value that our students are gaining from the 24/7 service. We noted topic patterns that matched up well with our curriculum (English and Speech are huge library users). We found that students are persistent advocates for themselves. A significant percentage logged back in, sometimes repeatedly. This is an important finding, and it would be interesting to see if it is replicated at other schools.

Are there ongoing technical barriers to student's use of virtual reference? Particularly when dealing with a commuter

school, do we understand what barriers students face when using virtual reference services? If this study is repeated with more recent transcripts, focusing on whether technical barriers have persisted is essential. As colleges shift toward hybrid and fully online learning, how much do we know about the technological infrastructure that our students are relying upon? Is this a question that is specific to commuter schools?

Reflecting back to the study done on the complexity of questions coming through chat reference.<sup>18</sup> They reported a level of complexity that is not reflected in these transcripts. How complex are the questions being asked by community college students? Is there a major difference between how university students and community college students are using virtual reference? The next round of transcript analysis will examine a more recent batch of transcripts and look at complexity and technological stability.

There are some immediate outcomes to “close the loop” on this project. Noting that it is lightly used, we have been doing more marketing. We review transcripts constantly, not waiting for transcripts to be referred for follow up. Within the QuestionPoint community, it is hoped that this study will open up a conversation. Let’s keep talking about instruction versus information. There are many common concerns between virtual and face-to-face reference practice. Let’s open up a conversation.

---

## CREDITS

This project is part of the program “Assessment in Action: Academic Libraries and Student Success” which is undertaken by the Association of College and Research Libraries (ACRL) in partnership with the Association for Institutional Research and the Association of Public and Land-grant Universities. The program, a cornerstone of ACRL’s Value of Academic Libraries initiative, is made possible by the Institute of Museum and Library Services.

## References

1. K. Brown, *Connect, Collaborate, and Communicate: A Report from the Value of Academic Libraries Summits* (Chicago: Association of College & Research Libraries, 2012), [http://www.ala.org/acrl/sites/ala.org/acrl/files/content/issues/value/val\\_summit.pdf](http://www.ala.org/acrl/sites/ala.org/acrl/files/content/issues/value/val_summit.pdf).
2. Lisa Peet, “Obama Launches Connect Home Initiative,” *Library Journal* 140, no. 13 (August 2015): 12–13. Education Source, EBSCOhost.
3. Sharon Naylor, Bruce Stoffel, and Sharon Van Der Laan, “Why Isn’t Our Chat Reference Used More? Finding of Focus Group Discussions with Undergraduate Students,” *Reference & User Services Quarterly* 47, no. 4 (Summer 2008): 342–54. Education Source, EBSCOhost.
4. *Ibid.*, 351.
5. V. Armann-Keown, C. A. Cooke, G. Matheson, “Digging Deeper into Virtual Reference Transcripts,” *Reference Services Review* 43, no. 4 (2015): 656–72.
6. New York City Department of Education Division of Disabilities and English Language Learning, *2013 Demographic Report*, [http://schools.nyc.gov/NR/ronlyres/FD5EB945-5C27-44F8-BE4B-E4C65D7176F8/0/2013DemographicReport\\_june2013\\_revised.pdf](http://schools.nyc.gov/NR/ronlyres/FD5EB945-5C27-44F8-BE4B-E4C65D7176F8/0/2013DemographicReport_june2013_revised.pdf).
7. Katherine M. Conway, “Exploring Persistence of Immigrant and Native Students in an Urban Community College,” *Review of Higher Education* 32, no. 3 (2009): 333, <https://muse.jhu.edu/>.
8. Lorri Mon and Joseph W. Janes, “The Thank You Study: User Feedback in E-mail Thank You Messages,” *Reference & User Services Quarterly* 46, no. 4 (Summer 2007): 58. Education Source, EBSCOhost.
9. David Ward, “Measuring the Completeness of Reference Transactions in Online Chats: Results of an Unobtrusive Study,” *Reference & User Services Quarterly* 44, no. 1 (Fall 2004): 46–56. Education Source, EBSCOhost.
10. Jennifer Waugh, “Formality in Chat Reference: Perceptions of 17- to 25-Year-Old University Students,” *Evidence Based Library and Information Practice* 8, no. 1 (March 14, 2013), 23, <https://ejournals.library.ualberta.ca/index.php/EBLIP/article/view/17911/14792>.
11. Naylor, Stoffel, and Van Der Laan, “Why Isn’t Our Chat Reference Used More?”
12. Vera Armann-Keown, Carol A. Cooke, and Gail Matheson, “Digging Deeper into Virtual Reference Transcripts,” *Reference Services Review* 43, no. 4 (2015): 664.
13. Jan H. Kemp, Carolyn L. Ellis, and Krisellen Maloney, “Standing By to Help: Transforming Online Reference with a Proactive Chat System,” *Journal of Academic Librarianship* 41, no. 6 (November 2015): 766. Business Source Complete, EBSCOhost.
14. Krisellen Maloney and Jan H. Kemp, “Changes in Reference Question Complexity Following the Implementation of a Proactive Chat System: Implications for Practice,” *College & Research Libraries* 76, no. 7 (November 2015): 959–74. Education Source, EBSCOhost.
15. Sarah Passonneau and Dan Coffey, “The Role of Synchronous Virtual Reference in Teaching and Learning: A Grounded Theory Analysis of Instant Messaging Transcripts,” *College & Research Libraries* 72, no. 3 (May 2011): 278. Education Source, EBSCOhost.
16. T. Brown, personal communication.
17. Bettina G. Hansel, “Question for Institutional Research,” email to Robin Brown, February 6, 2017.
18. Maloney and Kemp, “Changes in Reference Question Complexity.”