Using Content Analysis Software to Analyze Survey Comments

Bradford W. Dennis
Western Michigan University, brad.dennis@wmich.edu

Tim Bower
Keuka College, tbower@mail.keuka.edu

Follow this and additional works at: http://scholarworks.wmich.edu/library_pubs

Part of the Library and Information Science Commons

WMU ScholarWorks Citation
http://scholarworks.wmich.edu/library_pubs/24
Using Content Analysis Software to Analyze Survey Comments

Bradford W. Dennis and Tim Bower

Abstract

In order to get the most from LibQUAL+™ qualitative data, libraries must organize and classify the comments of their patrons. The challenge is to do this effectively and efficiently. This article illustrates how researchers at Western Michigan University Libraries utilized ATLAS.ti 5.0 to organize, classify, and consolidate the LibQUAL+™ comments.

Introduction

The intent of this article is to introduce into the literature a methodology that allows for the manipulation of qualitative responses to LIBQUAL+™. The LibQUAL+™ survey is an assembly of up to 27 closed-ended questions (22 core and five custom-local questions) and one open-ended question that asks respondents "Please enter any comments about library service in the box below." Many libraries have found these comments helpful since they complement the quantitative findings and provide a level of insight not captured by the 27 closed-ended questions. Flo Wilson of Vanderbilt University, for
instance, noted that "the qualitative information gleaned from these comments constitutes some of the most specific and valuable data gathered from the survey."¹

When the authors presented in LibQUAL™ panel sessions at ALA and ACRL in 2005, many librarians requested practical instructions on how to analyze the content of the comments from the open-ended question and how to use qualitative data analysis software to assist in performing the analysis. Although several researchers have written about the analysis of LibQUAL+™ quantitative data, very few have written about the analysis of qualitative data, in general, and even fewer have written about using content analysis software for analyzing open-ended comments. Klaus Krippendorff asserts that "content analysis provides new insights, increases a researcher's understanding of particular phenomena, or informs practical actions."² This article examines the recent studies that discuss analysis of LibQUAL+™ open-ended comments and how content analysis software was used in our analysis.

Review of the Literature

There are five recent articles that describe the qualitative analysis of LibQUAL+™ data—Julie Anna Guidry in 2002; Wendy Begay et al. in 2004; Lorraine Harricome and Bonna Boetcher in 2004; Amy Knapp in 2004; and Flo Wilson in 2004.³ In 2001, the LibQUAL+™ survey did not include the open-ended question box. However, many participants e-mailed comments to the LibQUAL+™ survey administrators at their institutions. Guidry described the content analysis of these unsolicited e-mail comments from respondents of the 2001 LibQUAL+™ survey at Texas A&M University. Guidry's
stated purpose was "to complement the quantitative findings of the spring 2001 LibQUAL+™ survey to refine the instrument…[and to]…introduce readers to the qualitative data analysis program, ATLAS.ti, which was used to conduct the analysis."4 We used this practical article as a guide through the content analysis process using ATLAS.ti software.

Begay et al. performed research similar to that described in this article, but they used different content analysis software. They examined the issues that were of greatest concern to their customers. They needed to systematically organize comments and interpret them in a way that would aid the decision-making process of the Strategic Long Range Planning (SLRP) Team. Begay et al. used QSR's N6 to analyze 303 comments "to determine strategic direction at the organizational and individual team level by triangulating the qualitatively organized comments with the quantitative LibQUAL+™ data, as well as other sources of customer data."5 To date, this is the only article that specifically addresses the analysis of LibQUAL+™ survey comments using content analysis software. It was also instrumental in planning this analysis.

Haricombe and Boettcher focused on the analysis of both quantitative and qualitative LibQUAL+™ data. The results were then used to plan and guide focus group interviews; this feedback was used in their strategic planning process. They used Microsoft Access to analyze 189 comments, but discussion of their content analysis process is limited.

Knapp discussed the use of both qualitative and quantitative data to improve both existing services and the restructuring of public services at the University of Pittsburgh. Although she states that the comments from the survey provide "some of the most
detailed information," she does not indicate how many comments were analyzed, the software used, or how she utilized content analysis to analyze the comments.6

Wilson, in her article describing the 2002 LibQUAL+™ survey at Vanderbilt University, also focused more attention on the analysis of the quantitative data rather than the qualitative data. At Vanderbilt, 380 respondents chose to leave comments. Although it is clear that Wilson identified areas in need of improvement and has begun the process of addressing them, she did not indicate if content analysis software were used to analyze the comments.

The Data, Method, and Software

In addition to the quantitative data from the closed-ended questions, a library that administers LibQUAL+™ receives a dataset of patron comments. Currently, LibQUAL+™ does not provide an analysis of the comments. Organizing and summarizing these open-ended comments adds a level of "rigor, breadth, complexity, richness, and depth that is not available by means of the quantitative analysis alone."7 Content analysis assists in systematically exploring the content of the comments and enhances the understanding gained through the quantitative analysis. At Western Michigan University (WMU), it provided an efficient way to organize 754 comments so that library employees could read comments relevant to their position and department or service area.8 Through content analysis, a more in-depth understanding of service quality as perceived by library patrons can be obtained and communicated than from just the quantitative data alone.
Historically, conducting content analysis—coding and categorizing text in particular—has been completed by hand and required a great deal of time. Today, however, there is the technology to "process great volumes of text in a very short time."9 Content analysis software helps researchers categorize and catalog comments into themes. However, this process of categorizing comments into themes is not automatic; the researcher must identify the themes or units. For content analysis presented in this article, units were identified within the comments by categorical distinctions.

"Categorical distinctions define units [themes] by their membership in a class or category—by their having something in common."10 The process of unit identification is an iterative process; that is, the researcher develops, merges, and collapses units while progressing through the analysis.

Software Selection

The authors used the ATLAS.ti content analysis software for all of the reasons that Guidry described in her article: "It provides quick access to the quotations of a particular code; it provides search features that let the researcher find patterns or strings within the primary document; it allows the researcher to assign more than one code to a quotation; and it allows the researcher to assign more than one quotation to the same piece of text."11 ATLAS.ti is flexible and easy to use. It is simple to make editorial changes; the researcher can rename, delete, merge, or unlink codes or quotations without difficulty. Quotations can be accessed quickly from almost anywhere in the program, and it is easy to create a variety of reports.
Another factor in the selection of this software was that ATLAS.ti was used by Colleen Cook, Fred Heath, and Bruce Thompson to organize, classify, and analyze interview transcripts to develop the theory behind LibQUAL+™. Thus, it seemed in-line with the principles under which LibQUAL+™ was developed. Furthermore, since LibQUAL+™ offered workshops on the use of ATLAS.ti, the authors reasoned that it would be possible to obtain help from LibQUAL+™ if necessary.

Although Atlas.ti proved a simple and satisfactory system, it is not the only software option available. N6, an alternative content analysis software system, provides similar features according to its Web site. The N6 software was selected by Begay et al. since a member of their team was already familiar with that product. Other libraries have used Microsoft Access or Excel to organize their LibQUAL+™ comments.

Coding—Keeping it Simple

Although each researcher will make decisions on the level of specificity of codes, a primary goal of content analysis is to categorize and catalog themes within the data. Most content analysis software includes several coding options. The most common are: open coding, code in vivo, and code by list. Open coding is used when the researcher creates a code for the first time; code in vivo is used when the research wants to use a direct quote as the code; and code by list is used when the researcher wants to use a code that has been previously used. Begay et al. indicated that their working group used the open coding method advocated by Anselm Strauss and Juliet Corbin.¹²
The group began by analyzing a small sample of the comments sentence-by-sentence and labeled the activities, actions and events described in those responses. These were coded as concepts. Similar concepts were, in turn, grouped together into larger categories. This work on an initial set of concepts and categories provided training in coding, and the resulting categories served as an iterative guide for continuing coding.13

Many of the codes used in this project reflect the central theme of the quotation, using a respondent's own terminology whenever possible. For example, the quote "There could be more designated areas for group study and more enforcement of people who are noisy" was coded as "group study space— increase" and "noise—enforcement." The ability to assign multiple codes to a single statement, as in this case, was essential for proper classification of customer comments.

The use of dichotomies, such as positive/negative, increase/decrease, easy/difficult, and comfortable/uncomfortable, were very useful in developing categories. Identifiers of particular services were also helpful in the development of codes and categories.

It was invaluable to develop a hierarchical code structure from general to specific. Although many comments were specific, general comments were most common. General comments, such as "I love the library" and "We need a better library," were coded as "library—positive" and "library—negative,” respectively. However, without specific indicators, these quotes reveal very little. The statement "I think you have a marvelous collection; I've always been able to find what I need" was coded as "collection—positive"
because the respondent refers only generally to the collection but unfortunately does not state which part (whether monographs, journals, databases, print, or online collection) was deemed "marvelous." If the comment or quotation were more specific, more specific codes were used.

After several codes were established, the "code by list" feature was utilized in order to keep the number of new codes to a minimum. The coding process proceeded faster once a relatively comprehensive set of codes was in place. The ATLAS.ti options for renaming, deleting, and unlinking codes were conducive to the iterative process of coding, as was the "undo" feature.

By the time the survey had been coded once, the authors had developed a comprehensive set of codes that were then used to recode the comments. Recoding provides a means to check that all comments are coded properly and assures that some degree of consistency and reliability has been maintained throughout the iterative coding process. Any coding discrepancies identified in the recoding process were resolved.\textsuperscript{14}

Codes with similar themes were examined and collapsed. Codes with few instances were merged with others of the same thematic nature. For example, there was a code for "books—misshelved" (2 quotations), a code for "books—misplaced" (2 quotations), and a code for "shelving error" (1 quotation). All three codes were collapsed into one, "books—misshelved" (5 quotations). This iterative process is at the heart of content analysis.

Code Frequency List
Once the iterative coding process was well underway, a list of all codes along with the frequency that each was used was created. The code frequency list is a very useful way to display and report findings, especially when presenting results to colleagues and administration. From the code frequency list, it was possible to identify the thematic categories that needed to be merged or collapsed, as well as to begin to learn about our customers' preferences and concerns.

The finalized code frequency list used by WMU Libraries is presented in table 1 (codes with frequencies under four are removed from this figure in order to save space). Although this coding was developed iteratively from the WMU LibQUAL+™ comments, these, or similar codes, are likely to have general applicability to many academic libraries because they are based on typical academic library functions and units. As such, they may provide a useful basis for other libraries as they begin their content analysis. Indeed, the code frequency list handout was the most often requested item by other librarians who attended the authors' ALA presentations.15

<table>
<thead>
<tr>
<th>Table 1. Code Frequency Table</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2004 LibQual + Codes – Alphabetical Order</strong></td>
</tr>
<tr>
<td><strong>Frequency Count – 754 Comments</strong></td>
</tr>
<tr>
<td>Codes within ATLAS.ti</td>
</tr>
<tr>
<td>Access - negative</td>
</tr>
<tr>
<td>Access - positive</td>
</tr>
<tr>
<td>Books - lost</td>
</tr>
<tr>
<td>Books - misshelved</td>
</tr>
<tr>
<td>Books - negative</td>
</tr>
<tr>
<td>Cafe</td>
</tr>
<tr>
<td>Catalog - negative</td>
</tr>
<tr>
<td>Cell phones - negative</td>
</tr>
<tr>
<td>Central reference - positive</td>
</tr>
<tr>
<td>Checkout - increase</td>
</tr>
<tr>
<td>Circulation - negative</td>
</tr>
</tbody>
</table>
The code frequency list showed that the majority of the libraries respondents' comments were positive and suggest that, overall, our customers "feel good" about their library. For example, the shaded areas in table 1 show "library—positive," 92, versus "library—negative," 10 (9:1 ratio); and "employees—positive," 96, versus "employees—
negative," 47 (2:1 ratio). However, some of the results appeared contradictory such as "collection—positive," 9, versus "collection—negative," 18 (1:2 ratio) compared to "resources—positive," 18, versus "resources—negative," 5 (3:1 ratio). In such cases, focus group interviews can help clarify how the library customers define the terms "resources" and "collection."

When interpreting output, it was important to keep in mind the changes that had taken place at the university and the library both during and preceding data collection. For example, library hours were reduced for a period prior to the administration of LibQUAL+™ but had been restored two months before LibQUAL+™ began. Nonetheless, the dissatisfaction over the reduction in hours still showed up as a major theme in the comments, and several respondents were obviously unaware that the hours had been restored. Similarly, the libraries' new Web site was made operational in January 2004, and we began the LibQUAL+™ survey in March 2004. Yet, it was unclear whether respondents were referring to the old Web site or new Web site in their comments about the site, especially in light of the unawareness of the change in library hours over a similar time frame.

It is rarely possible or even desirable to mandate total stasis in order to eliminate such uncertainties in LibQUAL+™; however, it is imperative to interpret findings from both qualitative and quantitative analyses in context. Happenings and changes that take place prior to and during data collection influence patron perceptions of services, but this can be informative and even helpful. In this case, the 2004 LibQUAL+™ survey provided data that clearly showed that our customers were strongly opposed to reducing library hours and, indeed, desired that hours should not just be restored but increased.
Although the reduction in hours created a crisis for a brief period of time, the LibQUAL+™ comments affirmed that the library is viewed as a critical institution on campus and clearly demonstrated to the university administration that our faculty and students viewed the library as mission critical and were willing to fight for it.

Categories and Subcategories

Each customer comment may include several quotations about a particular issue or may discuss multiple services that the library provides or collections that the library contains. This means that it is not unusual to find several codes per comment. After collapsing or merging codes with similar meaning, the ATLAS.ti networking feature was used to link related codes in order to build a graphical representation of the respondents' comments. This helped reveal relationships between and among comments. This feature was also helpful in organizing a large number of codes into categories and subcategories of related links.
Ultimately, five large categories were created (see figure 1). Three reflect the LibQUAL+™ domains: services (affect of service), environment (library as place), and resources and access (access to information). A category, "departments" was added for quotations that specifically identify a given library department. For example the statement "I love the Education Library" was coded "Education Library—positive." The final category is a miscellaneous one labeled "other." It includes codes such as "suggestions" for very specific recommendations and "misconceptions" about library
collections or services. "Misconceptions" provided issues to be addressed in the libraries' marketing endeavors and on our Web pages.

Figure 2 below displays only one of the five large categories that was identified because the entire network view is nearly five times the size of the figure presented. Figure 2 shows the "environment" domain (also known as the LibQUAL+™ library as place domain). The most general of comments are arranged on the left, progressing to the most specific on the right. Note how all of the codes or boxes (ATLAS.ti uses the term "nodes") on the right have at least one corresponding link to the left. In figure 2, the code (or node) "environment" on the far left directly links to the code "library" in the center of the network shown in figure 1. "Library" is the most general link of the entire network view. The code "computers—positive" is arranged at the far right of figure 2 because it is a very specific code. It is arranged under the codes "library" (the most general), then (from left to right on figure 2) under "environment," then "facilities," then "equipment," and then "computers."
Using the Data

Reporting to Colleagues and Administrators: Completing the Feedback Loop

One of the most important steps in moving from data collection and analysis to using findings for improvements is completing the assessment feedback loop. Shelley Phipps stated, "Staff must be involved in learning about customer needs, current dissatisfactions with the whole library, and future customer priorities as part of strategic planning." Reporting results to all of our colleagues (all library employees) was considered a crucial
step in the process of learning about patrons' perceptions of service quality and improving service for our customers.\textsuperscript{17}

The WMU LibQUAL+™ Task Force reported findings to the dean of libraries, all employees, and various groups responsible for specific functions. Each employee from every department in the library was able to examine all the quotations and also to focus on those that specifically applied to his or her areas of responsibility. To facilitate this, we organized the comment files by code (or category) in a file on the libraries' shared drive so that employees could readily sort through the code list to find those quotations applicable to their particular responsibilities and then read the relevant full quotations. For example, if someone were looking for all comments relevant to customer service on the shared drive, a report would be selected for "customer service—positive" and "customer service—negative." In addition, if a strategic stakeholder like the dean of libraries wanted the LibQUAL+™ Task Force to retrieve all examples of good customer service at the main library, the team could easily query the comment file using ATLAS.ti and create a report. The LibQUAL+™ Task Force made it known that they were willing to run queries on any of the codes or categories for any colleague, in addition to what was made available on the shared drive.

Knapp reinforces the idea that involvement of all staff with the comments is essential. She stresses that employees must be made aware that the data collected are our users' perception of the quality of library service. Knapp says,

\begin{quote}
When sharing the results internally, one of the first hurdles we as a library system had to address was how to approach the findings in a positive,
proactive way, instead of defensively. It was important for us to communicate to our staff that you cannot argue with perceptions—whatever they are, they are. Our challenge was to take these perceptions and attempt to design methods and services to address them.18

Using Qualitative Data to Support Quantitative Data

Content analysis was conducted in order to add a level of depth and breadth to the understanding of the service quality at WMU that is not available via the quantitative analysis alone. All too often decisions are based on anecdotal evidence or the opinions of what library personnel think customers want and need. The ability to base planning decisions on reliable quantitative data and robust comments from our customers is essential for improving quality of service. For example, certain unique quotations were coded as "reasons that I don't use the library" to prompt library improvement. This is valuable information and usually difficult to obtain, except through an anonymous survey. Libraries desperately need to know such information in order to address problems with service, the collection, or the library building and environment. The use of findings from quotes and this content analysis made the quantitative findings more tangible for those with limited statistical training.

Like most academic libraries, our customers identified the domain access to information as the most important. When examining the libraries' quantitative LibQUAL+™ data, the authors found that minimum acceptable and desired scores were higher in this domain than in the domains affect of service and library as place. The
quantitative data are supported by the qualitative data. More negative comments were made by customers regarding access to information than any other domain (see table 2). Table 2 includes the number of general and departmental comments. General comments are those comments that do not identify any specific domain, for example—"The library is great" or "I do not use the library." Departmental comments include those comments in which the library department, unit, or branch is identified, for example—"I love the Education Library" or "The Interlibrary Loan Department has helped me on several occasions."

Table 2 – The Number of Comments Table

<table>
<thead>
<tr>
<th>Number of Comments, Positive, Negative, Neutral and Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain or Comment Type</td>
</tr>
<tr>
<td>General</td>
</tr>
<tr>
<td>Library as place</td>
</tr>
<tr>
<td>Affect of service</td>
</tr>
<tr>
<td>Access of information</td>
</tr>
<tr>
<td>Departmental</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

The libraries have made many improvements in an effort to overcome negative perceptions in the access to information domain. Many faculty and students left comments indicating that they would like the collection improved. In particular, they would like an increase in the availability of full-text online articles. However, other comments indicated that some of our constituents were unaware of resources that the libraries had. Several comments also indicated dissatisfaction with remote access. To address both these issues, EZ Proxy was installed to simplify the log in process off campus, and the library Web site was improved.
The lack of awareness of services or resources may indicate a need to publicize them. In order to identify such needs, a code for "misconceptions" was created and used to help identify services that need promotion. For example, a few respondents suggested that the libraries offer some kind of online renewal service when, in fact, this option has been available for several years. Such findings were included in the internal report, and it was suggested that the strategic plan include a provision for increasing the awareness of the services and resources that the libraries offer. Informing the libraries' customers of what has been done reminds them that we are listening and valuing their suggestions. The Marketing Committee has been able to create ads such as table tent signs to be placed on all tables in all university library branches to illustrate the changes that have been made (see figure 3).
In the library as place domain, patrons indicated that they value a bookstore-like environment and are interested in a café located in the library, bringing covered drinks into the library, comfortable furniture, and an inviting environment. Similarly, customers stated that they also desire a quiet atmosphere (at least in designated areas), as is evident in comments about noise, noise reduction, noise reduction enforcement, and limiting the
use of cell phones. Such rich data give a better picture of what is needed to improve the library as place than would quantitative results alone.

The authors’ quantitative analysis had indicated that each user group (undergraduates, graduates, and faculty) within the College of Education was more critical of the Education Library (built in 1965) than they were of the main library (completely renovated in 1991). A stronger case for renovation of the Education Library was possible due to the ability to draw on comments about the library to support and reaffirm the SPSS (Statistical Package for the Social Sciences) analysis that had been done. A graduate student, for example, stated, "The facility in Sangren [Education Library] could be updated and made more physically comfortable, but the services are excellent." This and other quotes were essential in conveying the "state of library services" to stakeholders.

Conclusion

Shelly Phipps noted, "Research library organizations must design internal systems that help staff keep current with customer needs, understand the real causes for dissatisfaction, discover what would increase satisfaction, and focus staff efforts on improving services and creating new products." LibQUAL+™ brings us one step closer to this goal; and, by effectively utilizing qualitative data analysis software, local LibQUAL+™ administrators can quickly organize, classify, and consolidate LibQUAL+™ comments. The authors have analyzed patron comments about WMU Libraries in order to provide feedback to library administrators, departments, and
employees. The content analysis provided a level of richness that was essential for supporting and strengthening the findings from the detailed SPSS analysis of the quantitative data from the LibQUAL+™ survey. The analyses were both invaluable and highly complementary in the libraries' efforts to evaluate the quality of service our customers receive.

The findings from our analyses have been used in the library assessment plan, strategic planning, and strategic marketing. Content analysis of comments can help a library to learn more about its customers, to lobby for improvements, to report to colleagues and administration, and to market the library to strategic stakeholders.

The analysis of our 2004 qualitative comments and the resultant categories and subcategories will help the libraries analyze comments of our 2007 survey. There now exists an extensive list of categories that will provide a starting point for coding and comparison between the survey years. The authors believe that Knapp, indeed, is correct when she stated, "Perhaps one of the most significant benefits we as a library system derived from the LibQUAL+™ process was the increased commitment to the ongoing solicitation of user input and an awareness of the role of users in guiding public service initiatives."\(^{21}\)

Bradford W. Dennis is education librarian, University Libraries, Western Michigan University, Kalamazoo, MI; he may be contacted via e-mail at: brad.dennis@wmich.edu.
Tim Bower is assistant professor, applied and comparative sociology, Division of Basic and Applied Social Sciences, Keuka College, Keuka Park, NY; he may be contacted via e-mail at: tbower@mail.keuka.edu.

Notes


4 Guidry, 100.

5 Begay et al., 112.
6 Knapp, 160.


8 Seven hundred fifty-four of the 1,678 total respondents left comments.

9 Krippendorff, 16.

10 Ibid., 105.


13 Begay et al., 113.

14 The recoding of comments should ideally be done by the original researcher and by another analyst independently. Results of each analysis are then compared to establish inter-coder reliability. However, if no one else is available to perform the recoding, then it has to be performed solely by the original researcher.


17 Although we believe that every employee should have access to most of the reports, one file was confidential at WMU because it contained negative comments in which employees were mentioned by name. This file was given to the dean of libraries, who
then met with the named employees and their supervisors to discuss the matters of concern in confidence.

18 Knapp, 162.


20 Phipps, 643.

21 Knapp, 170.

This is a post print paper originally published as: