City University of New York (CUNY)

CUNY Academic Works

2014

Can't Get No Satisfaction: Lessons Learned while Developing an Assessment of Faculty Awareness and Attitudes towards Library Services

John Pell
CUNY Hunter College

Sarah Laleman Ward
CUNY Hunter College

Margaret Bausman
CUNY Hunter College

[How does access to this work benefit you? Let us know!](#)

Can't Get No Satisfaction:
Lessons Learned while Developing an Assessment of
Faculty Awareness and Attitudes toward Library Services

John Pell, M.Ed., MLIS, Sarah Laleman Ward, MLIS, MA and Margaret Bausman, LCSW, MSLIS

Abstract

This paper expands upon the presenters' recently published article in the *New Review of Academic Librarianship* titled: "Beyond Satisfaction: Understanding and Promoting the Instructor Librarian Relationship" that reported on the progress and the findings of an action research project intended to advance the Hunter College Libraries services to instructional faculty. This companion piece provides descriptions of conceptual, methodological, and practical challenges that were not covered in the article but may be particularly germane to other librarian practitioners interested in pursuing locally relevant assessment of faculty services.

Inception and Approach

In the fall of 2011, the Hunter College administration asked the Libraries department to assess the Hunter faculty's level of satisfaction with library services and resources. Assessment of faculty satisfaction is an increasingly common benchmark of library quality. For example, the Primary Research Group's *Information Literacy Benchmarks, 2013* specifically inquires if the respondents provide an evaluation format by which faculty can express their satisfaction with library services. Of the 60 academic libraries that participated in this study, 31% responded in the affirmative. This is certainly in keeping with the ACRL 2011 *Standards for Librarians in Higher Education* that explicates principles and performance factors that includes on-going assessment of service and outreach to collaborators including faculty. Further, value of academic library projects such as the Value of the Academic Library Initiative, LibValue and Working Together all contain components that address collaboration with and service to instructional faculty.

Challenge: Assessment of satisfaction is not a straight forward procedure

There are conceptual challenges in identifying meaningful characteristics to target as assessment points and enormous variation in the subjective understanding of what constitutes satisfying library engagement. The Hunter Librarians involved with this project grappled with three areas of local concern:

1. Faculty needs, and thereby satisfaction, can vary widely depending upon field of study and stage of career;

2. Questions of general satisfaction do little to assess faculty knowledge, needs and use of the diverse array of resources and services;

3. Measures and expressions of satisfaction can overly focus on the library as a facility and a collection of information sources which excludes the instructional and embedded functions performed by most academic librarians.

Lesson Learned: Meeting both the administrative request and generating products with meaningful application to Hunter's instructional community propelled the group toward an action research approach. The employment of an action research approach in the library context such as ours is intended to produce data that are actionable and can be directly applied to program development and service delivery. Further, this action research project is intended to be enduring and iterative in order to assess the impact of intervention while continually generating outputs that illuminate trends and identify continued or new needs (Reason and Bradbury, 2008).

Using the action research approach, the group pursues multiple lines of inquiry that assess faculty awareness of library resources, their use of library resources in their research and instruction, and their perceived satisfaction with library services. The goals are to raise faculty awareness of library resources and to create organizational and programmatic change that enhances library service to faculty in both their instruction and research.

Method: Informal Information Groups

The group facilitated 2 quasi-qualitative informal information groups comprised of instructional faculty in the winter of 2012. The group sought to elicit material across six (6) themes and three (3) categories of response type.

Themes	Response Types
In-person library use	Behavioral (factual descriptions of specific experiences)
Online library use	
Use of non-Hunter College Libraries	Affective (internal response to the experience)
Location of research	
Information literacy	Speculative (what else might be done or provided?)
Overall work experience	

Challenge: The identification and recruitment of appropriate participants

Lessons Learned: The cohort was an invited cohort of instructors with solid engagement to the library. While this was not a statically valid sample, the selection targeted those motivated to provide a balance of positive, negative and constructive material. Personal invitations, thoughtful scheduling and refreshments (lunch and "tea") were noted components in successful recruitment.

Method: Faculty Awareness of Library Services (FALS) Survey

Among other things, the material generated from the information groups formed the basis for the development of the Faculty Awareness of Library Services (FALS) Survey with which the research group has now collected three cycles of data. The purpose of the FALS Survey is to assess:

- What faculty instructors _know_ about the HCLs resources and services;
- What HCLs resources and services faculty _use in their instruction;_
- What HCLs resources and services faculty _use in their research;_

- If there are differences in awareness, use and level of satisfaction of HCLs resources and services between pre-tenure and tenured faculty, faculty of varying rank or academic status.

The 2012 FALS Survey instrument is a questionnaire consisting of twenty-four (24) questions, seventeen (17) multiple choice and seven (7) open-ended.   The 2013 FALS survey added one (1) question which uses a 5-point Likert scale to directly elicit perceived level of satisfaction. The 2014 FALS survey collapsed the open-ended questions into drop-down or ranking formats and added three new open-ended questions.

The group encountered five (5) challenges in the design and management of the FALS Survey:

Challenge: Structure

The research group selected 3 structural strategies intended to enhance recruitment for and to support completion of the survey.   These strategies included brevity (completion under ten minutes), a broad freedom of response in which respondents were allowed to skip questions, and placing the demographic section last so that the information concerning the central research questions were answered prior to any onset of survey fatigue.

Lessons Learned: Once they started, the vast majority of respondents completed the FALS Survey which would suggest that these strategies were successful. However, the degree of freedom of response had unanticipated implications for data analysis which will be discussed in more detail below.

Challenge: Incentive

The group considered the potential impact, positive and negative, of attaching an incentive to the FALS Survey.   Concern that the use of an incentive might be held in higher regard than attentive response was ultimately outweighed by the need to maximize the volume of respondents requisite for meaningful data analysis. The group used a service provided by the SurveyMonkey software to allow respondents to opt into a raffle for a $50 Amazon gift card upon completion of the survey.

Lessons Learned: In 2012, the FALS survey generated 240 responses and in 2013 it generated 227 responds. In 2014, the incentive service was no longer available and the FALS survey generated 109 responses. This would suggest that an incentive attached to the FALS Survey has the potential to doubles the response rate.

Challenge: Survey Tool

The Hunter College Libraries maintains a mid-level subscription to the online survey tool, SurveyMonkey. The group opted to use this resource as it is a well-known and familiar resource to both researchers and respondents, maintains respondents' anonymity and provides basic data analysis tools. While the question formats are varied, there are limitations. Moreover, the loss of the option to provide an anonymous, third-party incentive appears to have negatively impacted the 2014 collection.

Lessons Learned: The use of commercial software package requires a broad degree of flexibility.

Challenge: Internal Review Board (IRB)

The process of completing the online training modules required by the IRB is time consuming. Likewise, the preparation of IRB materials and use of the IRB submission site requires attention and time.

Lesson Learned: Build into the work-flow ample time to complete training as required, to complete IRB materials and to gain proficiency with the IRB submission website. Contact and utilize the support of those in the IRB Office who are a very knowledgeable resource intended to be of assistance to researchers.

Challenge: Distribution

The distribution process and marketing are all vetted through IRB. Distribution of the FALS Survey is facilitated through Hunter's Office of Communications in an email blast to all faculty. Typically,

this is a smooth process. However, a very small miscommunication that caused the 2013 survey to be distributed to all Hunter employees, rather than only the Hunter faculty, constituted a deviation from the IRB-approved protocol. It was fortunate that this was discovered within the first day the survey was open. The group immediately closed the survey, filed an Unanticipated Event Report with IRB, destroyed the data, and upon IRB approval, distributed the survey a second time.

Lessons Learned: When relying on an entity outside of your research group to facilitate part of the research process, be very specific and assume nothing. Additionally, check your data collector frequently to assure that the software is running well and the collection is proceeding as anticipated.

Analysis: Open-ended Questions

Challenge: Coding

The 2012 and 2013 FALS surveys contained five open-ended questions that yielded a total of 1,667 responses. This group of three researchers each reviewed and categorically coded the responses at least three times. This represents a total of over 15,000 person-question-views.

The group conducted the coding and recoding with this approach in order to analyze patterns in the data with a greater degree of reliability and validity than might have been achieved had the task been assigned to a single researcher. In order to provide a quantitative description of the categories of responses, it was necessary to ensure that the coding was not the artifact of a single researcher's perception. Further, this approach took variation in perceptions of content and intent between researchers into account and reported only on categories of response in which there was a high level of agreement.

Lesson Learned: Open-ended questions take a lot of time and effort to analyze, but they can be a very valuable step in survey development.

Challenge: Parsing content from intent

A category where the researchers found a low level of agreement is "criticism." For example, for one of the open-ended questions the group attempted to code for both content and intent. When analyzed for agreement, there was far less variations in the perception of the content facet than in the perception of the intent facet. In fact, standard deviation for the intent facet of "complaint" is four times as large as for the content facet of "librarian." This suggests that the qualities the group sought to analyze under the intent facet were highly subjective in nature. This point is further reinforced in that there were only three coding options for intent as compared to nine coding options for content, thus there is a greater probability of random agreement on the intent facet due to the small number of coding options. Still, the group found less agreement on the intent facet than on the content facet. As a result, the 2014 FALS Survey was modified to ask directly for criticism and praise.

Lesson Learned: It can be very difficult to make reliable inferences about latent qualities like intent; structure surveys in a way that makes the latent quality very clear.

Challenge: Interrater reliability and agreement

The previous section refers to the method of using the size of a standard deviation for a code as an estimate of the level of agreement about the responses assigned to that code. This method was a compromise solution to the problem of describing the level of agreement on the categories assigned to the open ended questions.

Initially the group sought to quantify discrepancies in our coding by reporting an interrater reliability score but found a challenge in selecting and applying a method.   The researchers learned that their intuitive concept of agreement did not match the statistical methods. For example, one might think that if two out of the three coders agreed on a code for a response that there was a high level of agreement. Two out of three is 0.66, which sounds pretty solid. The group learned that methods like Fleiss' Kappa evaluate the level of agreement relative to all possible rater pairs (Fleiss, 1981). So if two

out of the three coders agree on a rating, it means that only one out of the three possible pairs are in agreement; this gives a less impressive score of 0.33. Fleiss' Kappa also controls for the amount of agreement that may occur due to chance.

While it would have been nice to incorporate that level of rigor into the report, the group also learned that the coding scheme was not prepared in way that would be amenable to analysis with Fleiss' Kappa: the researchers applied multiple ratings to each response but the Fleiss' Kappa equation is set up to use one rating per rater per response.

Rather than reformat our coded data, the researchers settled on average values for each code occurrence to approximate the consensus of the coding team; the group then estimated agreement by calculating a standard deviation for the number of occurrences for each code for each question. The group only reported on the areas that achieved high levels of agreement. While there was some disappointment in resorting to an improvised method to express interrater reliability, it enabled the group to provide quantitative information about the validity of the coding scheme and it contributed an important role in the process of survey development. The group went on to use the categories of responses that emerged from the open-ended questions to structure multiple-choice questions on the current version of the survey. The group expects that this development will make it easier to conduct quantitative analyses on future survey results.

Lesson Learned: Effective assessments do not have to be complicated; less rigorous methods can sometimes provide useful information.

Analysis: Closed Questions

Challenge: Question structure and survey design

When the group created the survey it was not done with an eye toward how the data was analyzed. The group simply wrote the questions we wanted answers to and went ahead with it. This meant more

work on the back end when it came to analysis because the questions were not structured in a way that led to easy analysis of the data without some extensive clean-up. As mentioned above, the survey consisted of a combination of categorical data (tenured/untenured, rank, etc.) and free text fields. When it comes to analysis, categories are much easier to deal with than free text.

Lesson Learned: Design better-structured survey questions at the outset. Ask yourself:

- What you hope to find in the data you are collecting and what might be the most effective way to gather and analyze that information?

- Can you force a choice through a drop-down menu or ranking question, or is it necessary to leave an open text box? Is an "additional comments" box sufficient?

Challenge: Statistical Analysis

A criticism sometimes leveled at research studies in our field is that researchers often do not make use of inferential statistics and instead rely solely on the presentation of descriptive statistics (Hightower & Scott, 2012; Robbins, Engel & Kulp, 2011). Descriptive statistics such as frequency counts and percentages provide a summary snapshot of the population of respondents but not much more than that. The researchers sought to utilize inferential analysis where and as applicable.

The data derived from the closed questions on the FALS survey is categorical, meaning that even if the data is in numerical format the numbers are not meaningful aside from the categories they represent, e.g. 1=Lecturer, 2=Instructor, 3=Assistant Professor, etc. To put it another way, one cannot calculate means, medians, and standard deviations of categorical scores. Therefore, categorical data eliminates the possibility of certain types of statistical tests called parametric tests which require meaningful numerical scores on interval or ratio scales.

However, categorical data can be analyzed beyond the simple reporting of descriptive statistics by utilizing a chi-square test. The chi-square test for independence evaluates the relationships between variables, so it is a correlation based on the proportions or frequencies of respondents in each category instead of a numerical score (Gravetter, 2007). The researchers intend to pursue even more sophisticated analysis, probably a logistic regression, since we have a longitudinal set now of three years of survey data.

Lesson Learned: There is probably a way to press beyond simple descriptive statistics to a deeper, more meaningful analysis of the data. Keep an eye out for it, or plan for it at the outset of your study. It will add value to your reporting and elevate the research in our field to a higher level.

Challenge: Preparing the data for analysis

Anyone familiar with data analysis knows that there is always a certain amount of cleaning up that needs to happen prior to analyzing the data. Decisions must be made regarding missing data and how that will affect the analysis and reporting. The group decided to include all the data available, acknowledging that there were discrepancies in the final numbers because respondents were allowed to skip questions. The only cases that were eliminated from the report were those who only answered the informed consent question and nothing else. It was the general assumption that these individual cases represented people primarily interested in the incentive rather than participating in the actual survey and it did not impact the data analysis.

Recoding is up to the individual doing the analysis since what others will ultimately see is the final product. However, the coding scheme should make logical sense and there should be a key available to identify the codes so that others can access and utilize the data as well.

Lesson Learned: Cleaning and coding takes lots of time. Think about what you want to do with the data before you jump in and plan ahead as much as possible.

Challenge: Presenting effective visualizations of data

Multiple bar charts representing one piece of information each are not very meaningful or visually interesting. A reviewer comment for our recently published paper said, "The paper can be improved by including meaningful charts/graphs. It appears that the researchers learned a great deal about their users but the charts do not reflect those findings." One of Edward Tufte's principles for the visual display of quantitative information states, "The same ink should often serve more than one graphical purpose...Mobilize every graphical element, perhaps several times over, to show the data." (p. 139). Keeping this principle in mind the group revised the graphics to include the most amount of information possible while still maintaining clarity. In consultation with a data analyst, we arrived at a simpler graphic, a single bubble chart, which represented three bar charts worth of data.

Lesson Learned It can be extremely helpful to get an outside viewpoint. Explaining what you are attempting to convey to someone unfamiliar with the research will help clarify your purpose as well as bring a fresh perspective to the work. And, keep Tufte in mind and try to "mobilize every graphical element."

<u>**Assessment Value**</u>

There are several aspects of this action research project that resonate with the principles of ACRL Standards for Higher Education - in particular the principle of libraries conducting "continuous planning and assessment to inform resource allocation and to meet their mission effectively and efficiently" and the principle of libraries engaging "the campus and broader community through multiple strategies in order to advocate, educate, and promote their value" (ACRL, 2011). While national

benchmarks may not be locally meaningful, the value of the action research approach lies in its ability to effectively unite these two principles on a continuum of assessment and programmatic action.

Moreover, the assessment efforts each contain elements of outreach. For example, the informational groups provided with face-to-encounters with our faculty users and created opportunities to promote new librarian/instructor relationships while the FALS survey has shown that it serves to raise awareness and promote communication with instructors.

Reciprocally, the outreach efforts developed in response to assessment findings have also created new assessment opportunities. For example, as a result of the 2012 FALS Survey, the Hunter Libraries began hosting a yearly new faculty orientation luncheon. Not only is this a chance for librarians to have conversations with new faculty about their needs and interests but also creates another opportunity to assess the efficacy and perception of library services to the faculty.

Lastly, this project has added value to the library through its capacity and team building functions which can serve to promote a cultural shift within the community away from the "librarian as procurer" model toward an embedded "liaison librarian as partner" model.

Going Forward

The research group will continue to use the current action research approach for the foreseeable future in order to develop a solid longitudinal dataset. Not only would such a dataset identify trends over time but also would support a broader application of statistical analysis such as logistical regression to test for causal relationships between the variables. Such endeavors would require an expansion of the research team to include a primary statistician.

Likewise, the research team will continue to assess and refine the FALS Survey itself. The addition of a statistician to the team would support a more rigorous approach to the collapsing of qualitative queries into categorical options by employing higher level analysis to test inter-rater reliability.

Moreover, the research group is exploring ways to employ mixed methodologies which could include qualitative methods such as focus groups and ethnographic methods such as critical incident interviewing.

Lastly, while the locus of the project is intentionally local at this time, in the coming years, the research group hopes to entertain the possibility of expanding the FALS Survey to other CUNY campuses. Not only might this be a service to our colleagues employed in similar efforts, but also may shed light on the efficacy of the overall survey design and the generalizability of its findings to other settings.

References

Association of College and Research Libraries. "Standards for Libraries in Higher Education." Accessed June 2, 2014. http://www.ala.org/acrl/standards/standardslibraries.

Creaser, Claire, and Valerie Spezi. "Working Together: Evolving Value for Academic Libraries." Ed. Loughborough University LISU. 2012. Accessed June 2, 2014. <http://libraryvalue.wordpress.com/>.

Fleiss, Joseph L. "The Measure of Interrater Agreement." In *Statistical Methods for Rates and Proportions.* 2nd Ed. New York: John Wiley & Sons, 1981. 212-236. Print.

Gravetter, Frederick J. "The Chi-Square Statistic: Tests for Goodness of Fit and Independence." In *Statistics for the Behavioral Sciences.* Belmont, CA: Wadsworth Thomson Learning, 2007. 580-609. Print.

Hightower, Christy, and Kerry Scott. "Infer More, Describe Less: More Powerful Survey Conclusions through Easy Inferential Tests," 2012.

Mays, Regina, Carol Tenopir, and Paula Kaufman. "Lib-Value: Measuring Value and Return on Investment of Academic Libraries." *Research Library Issues*. 271 (2010): 36-40.

Oakleaf, Megan. "The Value of Academic Libraries: A Comprehensive Research Review and Report for

the Association of College and Research Libraries." 2010. Accessed June 2, 2014.

<http://www.acrl.ala.org/value/?page_id=21>.

Primary Research Group. "Information Literacy Efforts Benchmarks, 2013 Edition." 2013. Accessed June

2, 2014. <http://www.primaryresearch.com/read_more.php?press_release_id=87>.

Reason, Peter and Hilary Bradbury. "Introduction." *The SAGE Handbook of Action Research: Participation*

Inquiry and Practice. Eds. Reason, Peter and Hilary Bradbury. 2nd ed. Los Angeles: SAGE, 2008.

Print.

Robbins, Sarah, Debra Engel, and Christina Kulp. "How Unique Are Our Users? Comparing Responses

Regarding the Information-Seeking Habits of Engineering Faculty." *College & Research Libraries*

72, no. 6 (2011): 515–32.

Tufte, Edward R. *The Visual Display of Quantitative Information*. 2nd Ed. Cheshire, CT: Graphics Press,

2001. Print.