Shopping for Sustainability: Re-Envisioning the Secret Shopper Assessment

Tricia Boucher and Jessica McClean
Texas State University, USA

Introduction
We at Alkek Library have used the Secret Shopper assessment for several years to measure services as well as to inform initiatives to improve customer service. However, the Secret Shopper assessment in its present form requires tremendous time and energy from staff and provides limited, and seemingly diminishing, returns. In summer 2018 we began to research and review Secret Shopper to see if we could discontinue this time-consuming assessment. Unfortunately, we found that it is necessary for both reporting on our outcomes and understanding our services better through multiple types of data. Since we must continue it, we are working to make the Secret Shopper assessment more sustainable. Through research, examining our own experience, and understanding data collection in Alkek Library, we are developing a more strategic and sustainable assessment using economies of scale and an outcomes-based approach. A more sustainable Secret Shopper assessment will provide us with multiple types of data—descriptive, actionable data as well as measurements of success—about more types of user experiences, while also managing the assessment with less staff time commitment and effort.

Research
For Alkek Library’s Annual Report, we passively collect huge amounts of data that describe a high-level overview of services, so we expected that this would be enough to replace the Secret Shopper assessment. Unfortunately, descriptive statistical data does not give us the full story or the information we need to improve. Outcomes-based, mixed-method research—using a deliberate mix of quantitative and qualitative assessments to address a specific question or issue—provides detailed information on complex and multifaceted service transactions.1 Mixed-method research can ensure useful data are collected, check for consistency of findings generated by different sources, and provide actionable information.2 Using both qualitative and quantitative methods to measure service interactions fleshes out a larger story of services provided, giving them color and detail, and improves the quality of all data collected.3

However, incorporating qualitative methods into the assessment mix to better understand the user experience of services is notoriously difficult. User surveys are one option, but these put the onus on the user to provide information. And in the case of service points where patrons move on after they complete their transaction, a survey would be either ignored or hurriedly completed without the introspection needed to provide usable qualitative information. Waiting to send a follow-up survey decreases the likelihood of reliable responses.

Another option for collecting qualitative data about service points is direct observation. Direct observation of services does an excellent job of providing information missing from surveys and comments, but it is also difficult to implement.4 We have not employed this method at Alkek because, while the burden is not on the users, direct observation can be intrusive, takes a large amount of staff time and effort, and potentially infringes on the rights of users and staff.5

The Secret Shopper assessment is a third option for collecting qualitative data on service transactions. Borrowed from the retail and hospitality sectors, Secret Shopping is a technique for evaluating services in which trained evaluators work “undercover,” asking prescribed questions at service points and then immediately recording details of their experience using a reporting document.6 Secret Shopping is unique among service assessments in that it gathers information from natural transactions without directly putting a burden on either the staff providing the service or the general user.7 It is an excellent replacement for third-person qualitative observation, with the added benefit of presenting the service from a user’s perspective. It is also flexible enough to collect quantitative data using the same reporting tool.8 This is useful as service

742
points often require complex and multifaceted behavior on the part of the staff, which a mixed-method approach can capture well.9

However, there is some disagreement in the literature about what the Secret Shopper assessment can measure. While some libraries used the evaluation to measure user satisfaction, the majority did not. Their reasons for not doing so included: (1) examining user experience is not the same as examining user satisfaction; (2) even with good customer service, the correct information (or process, technology, etc.) may not be satisfying to the user; and (3) the evaluation can only measure what the shopper, as an observer, can describe or measure, including the actions and behavior of the staff, the accuracy of the answers received, and their own perceptions of the encounter.10

Despite this disagreement about the limits of the assessment, most libraries use Secret Shopper for a similar purpose: to explore known or suspected service issues, to metaphorically hold a mirror up to the staff, and to assess specific outcomes. These outcomes include practicing continuous quality management and proactive improvement; measuring employee behaviors and “soft skills;” measuring procedural and answer accuracy; and measuring space attributes.11 Some have used the assessment in different ways, to successfully measure and describe physical spaces and technological interactions.12 Others have even used it to examine wayfinding by writing broader questions and allowing shoppers to choose and locate the service point they thought would best be able to answer their question.13

Secret Shopping in Alkek
Alkek Library’s Secret Shopper assessment was envisioned in 2014 by Sarah Naper, director of Research and Learning Services (RLS), as a method of collecting new data about the library’s service points. Earlier in the year, the library had restructured reference services, transitioning to a triage model and training student workers to staff the reference desk. The pilot assessment went ahead in fall 2014 and covered all points of patron-staff interaction within RLS: circulation/reserve desk, reference desk, periodicals/media desk, government information desk, interlibrary loan office, online chat, and stacks management.

Three library staff members served on the team that designed and implemented the pilot assessment. Team members wrote question prompts that pulled from statistics, chat transcripts, and personal experience and then developed them into more complete “scenarios.” Scenarios consist of information about the service point to be assessed, a question to ask, background details to flesh out a believable transaction, and information about the resources a staff member might use to answer the question. Some scenarios anticipated a referral to another desk or department, so shoppers were told that they were not required to follow through on a referral; they were asked to focus on their experience at the initial point of contact, not necessarily on the completeness or accuracy of the answer. Finally, shoppers were asked to record their responses in a secure, online form. Shoppers were recruited from the university’s Alpha Phi Omega chapter and were each responsible for two scenarios. The shopping period lasted two weeks, and shoppers could ask their questions at their own convenience during regular service hours.

Following the shopping period, the team anonymized the collected data; shoppers’ names were removed from their responses, and specific details about the interactions (e.g., time of day, identifying information about staff members) were scrubbed and replaced by neutral pronouns and descriptors. Supervisors received an overview report that summarized all responses and detailed reports on their respective service points.

After eight semesters of running the assessment, we have made minimal changes to it besides modifying and adding scenarios to accommodate changes to library services. Through fall 2018, Secret Shopper was run each semester by the user experience librarian and one other library staff member. After the first semester, Alpha Phi Omega could not provide enough student volunteers, so student workers from other areas of the library—including technical services and special collections—were recruited as shoppers.

Initially, supervisors had a generally negative impression of the Secret Shopper assessment. To alleviate these misgivings, supervisors were asked for input on which skills should be tested and were assured that no
individuals would be singled out and that the results would not be shared widely. As Secret Shopper has been repeated, supervisors have embraced the assessment more. Because the assessment has been performed consistently over an extended period, repeated issues are eye-opening, and isolated negative interactions stand out more. They have also appreciated opportunities to celebrate positive responses.

Although supervisors have never been given a formal charge to put Secret Shopper results to use, several training initiatives have come out of the assessment. Most notably, student worker training across the department has been expanded to include a weekly newsletter and a day of training activities before each fall semester. Training initiatives focus particularly on consistency of information across service points and reinforcement of best practices and policies. Permanent departmental staff also now receive a similar weekly update email.

Over time, however, we have begun to see diminishing returns from the Secret Shopper assessment. While our time and energy invested—scheduling multiple training sessions to accommodate a small number of shoppers and following up with them to ensure completion—has not decreased over time, the quality of our results has. Some of this decline is because our shoppers are also library student workers. This work is not part of their regular job duties, and we have difficulty incentivizing high-quality responses. The student shoppers have noted that critically assessing colleagues is uncomfortable at best and have admitted that they tend to be less critical in case they are also being assessed. We who run the assessment share the blame for diminishing returns. Our reporting document effectively measures our reported outcomes but has not been updated or expanded. We need to continuously update our reporting document to provide critical and actionable information to promote improvement.

Current Project
Because of the time and effort involved in running the assessment and our declining return on investment, we would have liked nothing better than to sunset our Secret Shopper program. However, library administration intends to continue using some of the quantitative data collected in Secret Shopper as reported outcomes to accrediting bodies. After quickly reviewing other data we passively collect, we were not able to find another source of data to meet that required purpose, nor do we have another tool that could satisfactorily collect it. As it cannot be easily replaced, we cannot discontinue Secret Shopper.

To make Secret Shopper truly sustainable we need a greater return on investment from our labor and the labor of our shoppers, so we plan to overhaul and expand the program. Economies of scale will allow us to assess more outcomes than before and fine-tune the collection methods to provide more accurate and tailored data. This requires that we clarify our outcomes and expand the charge to use the Secret Shopper assessment to its fullest potential.

We expect two outcomes from this process: simplification of staffing and a gain in both quantity and types of data that will allow us to successfully measure and improve outcomes. We gain this data not only by increasing the number of scenarios run at each service point, but also by adding non-service-point scenarios, expanding scenario questions, and expanding the reporting document to collect targeted, actionable data where necessary. In this way, we can use the Secret Shopper assessment to its greatest potential: to gather data that can help us make changes and improvements to both our service points and other user experiences within the library.

Simplify Staffing
In our research we found multiple models for recruiting shoppers, including outsourcing, working with business faculty, and using the Human Resources department or existing quality assurance teams. Partnering with faculty fits well with our subject liaison program, so we have decided to use the Secret Shopper assessment as a chance to liaise with business faculty. We are currently working to partner with at least one business faculty to make the Secret Shopper assessment a class assignment. This will provide more shoppers, a centralized training place and time, and a built-in incentive for students to complete the assessment and provide high-quality responses.
More Data/Actionable Data
As we noted above, expanding the Secret Shopper assessment to bring in more data, as well as more actionable data, will give us more bang for our buck. However, since we do not want to duplicate other data collection currently being done, there are multiple steps to this process:

1. Run a data scan;
2. Run an outcomes scan and choose outcomes;
3. Analyze data against outcomes to choose data and data types needed to provide a mixed method approach to specific outcomes;
4. Modify scenarios and reporting documents to bring back the appropriate data.

Data Scan
Our first step was to review the Secret Shopper assessment to see how we currently use it for reporting outcomes and what data it provides for improvement. Our reporting document provides excellent information for the service points it covers, delivering data that both measure and give information for improvement (see Appendix). However, it only measures service points and does not cover patron-library transactions that do not include staff.

We then reviewed other collected data. Alkek Library staff collect both annual report data and outcomes-based data. The more passively collected annual report data is mostly quantitative, creating descriptive statistics that give administration a high-level overview of library uses and practices throughout the year: headcounts, classes taught, etc. The qualitative data we collect, such as chat transcripts, are not easily analyzed. Because these data are unrelated to clear outcomes, they require more analysis when queried for decision-making purposes. The more actively collected, outcomes-based data is collected to meet goals and explore issues. These generally come from targeted surveys or other instruments and are more easily analyzed to answer questions or support changes. We then gathered all the data we collected related to the user experience and labeled it by type (qualitative or quantitative) and collection method (survey, headcounts, chat transcripts, etc.). Organizing this information is necessary to avoid duplicating data when we modify the Secret Shopper reporting document.

Outcomes Scan
Outcomes drive data collection, so we began this step by gathering all library outcomes related to the user experience. For us, this involved reviewing outcomes reported to outside accrediting agencies, as well as internal outcomes related to our strategic plan and mission. We also included any data and data collection methods that measure the outcomes. In the future, we expect to find additional short-term, targeted outcomes from multiple sources, including in our library’s strategic plan, student complaints, journey maps, interactions with technology, library processes, space, and furniture. We are especially interested in using the Secret Shopper assessment to measure user experience outcomes that are not tied to a service point.

Another way we will find outcomes to measure is through staff. Like us, most libraries in the literature had librarians and library supervisors write the Secret Shopper scenarios and reporting document. However, all groups that run the Secret Shopper assessment also report that, in order to successfully implement improvement plans, staff buy-in and involvement is necessary. Here we found a strange dichotomy within the literature and in our own experience—for the most part, the extent of seeking staff buy-in is informing them that the assessment will take place and reassuring them that the results will not be used against them. This top-down model is not sustainable, as it requires time and effort to convince staff that the assessment is positive and any resulting improvement plans necessary. We believe that by using an organizational justice model that includes not only supervisor input but also staff and student worker involvement in the process of creating outcomes, we can create better buy-in for the assessment and make the results more meaningful.
We plan to work with both staff and student workers to find outcomes that are meaningful to them, then revise the scenarios and reporting document to incorporate them.

**Data Analysis**

We have created a process to plan data collection that is far messier than the graphic and list below suggest.

The process generally runs like this:

- Decide what we want to know (outcome), and what we will do with the information when we have it.
- Decide exactly what we want to assess to show whether we are meeting that outcome or not (e.g., are we measuring a specific behavior, answer accuracy, ease of use?).
- Decide what metrics will provide the correct data (e.g., do we need a Likert scale to show measured change over time, or a binary choice question to give direction for making a change?).
- Review the existing data to see if anything we currently collect will be appropriate.
- Find any gaps in the data collected to measure that outcome and decide if the Secret Shopper assessment is an appropriate instrument to collect it.

Using the outcomes as our starting point, we have begun to compare data we currently collect against desired outcomes to find the gaps that the Secret Shopper assessment could be used to fill. Our current Secret Shopper reporting document is not the only source of data we collect on service point interactions, but, as you can see, it is able to provide multiple types of data on one issue.

**Data Analysis Table 1: Using Secret Shopper for Multiple Types of Data**

<table>
<thead>
<tr>
<th>MEASURE OUTCOME: Improve staff-user interactions shown by</th>
<th>WHAT ASSESSED</th>
<th>METRIC</th>
<th>INSTRUMENT</th>
<th>DATA TYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvement over time</td>
<td>User perception of behavior</td>
<td>Likert scale for specific attributes</td>
<td>Secret Shopper</td>
<td>Quantitative—measure</td>
</tr>
<tr>
<td>Improvement over time</td>
<td>User perception of interaction</td>
<td>Likert scale for specific attributes</td>
<td>LibQUAL Lite</td>
<td>Quantitative—measure</td>
</tr>
<tr>
<td>Exploration of potential issues</td>
<td>User description of interaction</td>
<td>Free text</td>
<td>Secret Shopper</td>
<td>Qualitative—mine for specifics</td>
</tr>
<tr>
<td>Assess specific behavior</td>
<td>Were you acknowledged?</td>
<td>Binary yes/no</td>
<td>Secret Shopper</td>
<td>Quantitative—change action</td>
</tr>
</tbody>
</table>

We are in this process now, and we have already found that there are opportunities for the Secret Shopper assessment to fill in data gaps, particularly where outcomes have applicable quantitative data that reveal opportunities for improvement but not the qualitative data that would describe what could be improved. There are also opportunities in which qualitative data have shown specific items that could be improved and
a simple yes/no question about that item could measure whether or not there has been improvement. Here is an example of how to apply this to a different outcome.

### Data Analysis Table 2: Using Secret Shopper to Provide Mixed-Method Results

<table>
<thead>
<tr>
<th>OUTCOME: Improve user satisfaction with cleanliness shown by</th>
<th>WHAT ASSESSED</th>
<th>METRIC</th>
<th>INSTRUMENT</th>
<th>DATA TYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaction with cleanliness (over time)</td>
<td>User perception of general cleanliness</td>
<td>Likert scale for user satisfaction</td>
<td>Survey</td>
<td>Quantitative—measure</td>
</tr>
<tr>
<td>Satisfaction with cleanliness (over time)</td>
<td>User perception of restroom cleanliness</td>
<td>Likert scale for user satisfaction</td>
<td>Survey</td>
<td>Quantitative—measure</td>
</tr>
<tr>
<td>ADDED: Exploration of potential issues</td>
<td>User description of cleanliness</td>
<td>Free text</td>
<td>Survey AND Secret Shopper</td>
<td>Qualitative—mine for specifics</td>
</tr>
<tr>
<td>ADDED: Specific issue (derived from exploration)</td>
<td>Is there soap in the dispensers?</td>
<td>Binary yes/no</td>
<td>Secret Shopper</td>
<td>Qualitative—change action</td>
</tr>
</tbody>
</table>

Using multiple data types, we gain useful and actionable information to provide insights towards improvement. Because the date and time are included in the Secret Shopper assessment, we can see if there are trends specific to particular times or days of the week. We can also gather data on a specific desired attribute (in this case, full soap dispensers). This allows us to find problems, measure the solution, and potentially improve scores on the existing quantitative assessment. Essentially, it helps improve the user's experience. Over time, we expect to repeat this process with other outcomes.

**Modify Scenarios/Reporting Document**

The two major changes to the Secret Shopper assessment will be to modify the scenarios and the reporting document. The number of scenarios will expand to include non-service-point interactions with technology and spaces. These scenarios will need to be slightly different, as they may have to provide the shopper with more or less guidance in order to make the experience more realistic. The reporting document will also need to be changed to ensure data gaps are filled and bring back more information. This and the expansion to measuring non-human interactions with the library means that we may need to create different reporting documents for different types of scenarios.

We expect to pilot minor changes to existing scenarios and/or the reporting document in the spring 2019 Secret Shopper assessment. The larger scale changes we are looking to complete—translating multiple gaps in the data into changes in the scenarios and reporting document(s) and systematically tracking those changes—are part of a lengthier process that we expect to start this spring.

**Next Steps**

Since our scope enlarged dramatically, we are behind on our timeline. Our next step is to continue the process of data analysis. But prior to complete overhaul, we can start making changes to the Secret Shopper assessment immediately. We will choose one or two outcomes to pilot, create the scenario(s), and modify the reporting document to collect the necessary data. Once we have completed this step, we will work with business faculty to simultaneously run the Secret Shopper assessment using their students and our student workers. We will then assess their student responses against those from our student worker shoppers to see
if better data are collected. Once the results are satisfactory, we will replace our current Secret Shopper assessment with the piloted replacement, continue working with faculty to provide students for shopping, and expand the entire assessment further as our data analysis continues.

As we move through this process, we realize that the changes that could come from scaling up could be extremely beneficial to us. We expect that scaling up means that we will be able to do more of the higher-level work, more decision-making, and more analysis, which is less time-sensitive and more easily scheduled at less busy times of the semester. Additionally, it means that we will be able to systematize our work, including getting input from staff and reporting back to them; creating a system for tracking what is being explored and measured; updating scenarios and reporting documents; and leveraging partnerships with faculty to centralize training and remove the burden of shopper follow-up.

Where Alkek Library seems to be unique is that we repeat the Secret Shopper assessment every semester. We have maintained this schedule for several years, and we see this consistency as a strength, as we have gained experience and learned lessons over the years. It has helped us track performance over time and will help us track the effectiveness of improvement initiatives. This consistency will also help us evaluate the processes outlined in this paper.

Lessons Learned
Once our preliminary research showed that we could not get rid of the Secret Shopper assessment, we thought it would be a simple project—about a semester—to choose a few outcomes and update the reporting document to bring back data on those items. We will pilot changes doing just this, but in order to keep from duplicating data collection efforts, we decided it was important to begin with the data and outcomes scans and complete a comprehensive analysis of both. This larger scope, along with staffing shortages and the fact that we do not have a single person in charge of all library assessment, has dramatically lengthened our timeline. In short, fully exploiting the Secret Shopper assessment is more work than we had planned.

But as we have moved through this process, we have also found it to be an opportunity far beyond our initial plans. We are hoping that this expansion of our scope means that, when we are done, the Secret Shopper assessment will have become:

- An opportunity to partner with faculty and provide students with a real-world experience;
- A means by which to investigate multiple aspects of the user experience;
- A more powerful tool when used in conjunction with other data collection methods;
- Part of a larger outcomes-based data collection strategy that limits the duplication of effort and the number of assessments we run, and;
- A force for change while at the same time being simpler for us to manage.

Conclusion
Our current Secret Shopper assessment is charged with both measurement of and continuous improvement of service transactions—something we will continue to do—but we are excited to take this opportunity to expand the assessment to measure multiple aspects of the user experience. Doing so with a clear vision of how that data will work with other data currently collected should make the assessment a far more powerful tool, using mixed-method research to provide actionable information that both measures and provides insights as to how we might improve. Based on our experience, research, and our outcomes-based data analysis, over time we will be piloting multiple changes to our Secret Shopper program in order to fully exploit the assessment, and ultimately, to improve the user’s experience of the library.

—Copyright 2019 Tricia Boucher and Jessica McClean
Endnotes


Appendix: Secret Shopper Reporting Document

Start of Block: Default Question Block

Thanks for your willingness to be a secret shopper. Your feedback will help us provide better service to Alkek Library patrons.

1. What is your name? This will help us make sure that you are credited for your service. Names will be removed before results are reviewed.

________________________________________________________________

2. Which scenario are you evaluating?

3. When did you ask this question? (Please include date and time.)

________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________

4. How busy was it? What else was happening at the same time? (Please type NA for telephone or chat questions.)

________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________

5. What was the person doing before you approached them? (Please type NA for telephone or chat questions.)

________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________


6. Were you greeted when you approached the service point or asked for help?

- Yes (1)
- No (2)

7. Please respond on each line. The person who helped me...

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree (1)</th>
<th>Disagree (2)</th>
<th>Agree (3)</th>
<th>Strongly agree (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>was knowledgeable (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>was competent (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>was courteous (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>was friendly (4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>was positive (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>was attentive (6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>was patient (7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>was professional (8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>made me feel welcome (9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Do you have any comments about how the person interacted with you? (optional)

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

8. Please respond on each line. The person who helped me...
<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree (1)</th>
<th>Disagree (2)</th>
<th>Agree (3)</th>
<th>Strongly agree (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>took an appropriate amount of time. (1)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>asked follow-up questions. (2)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>provided a thorough answer. (3)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
</tbody>
</table>

Do you have any comments about how the person answered your question? (optional)
________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________

9. Did the person refer your question?
   
   o Yes (1)
   
   o No (2)
   
   o Not applicable (3)

If you answered “Yes,” where did the person refer you?
________________________________________________________________

10. Please describe the interaction that you had with the person that assisted you (almost a play-by-play). Include any information that you think might be helpful to us.
________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________
11. What would you have liked the person to do that they didn’t do?

________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________

12. Is there anything else that you want to share about your experience?

________________________________________________________________
________________________________________________________________
________________________________________________________________

Thank you so much for your help!

Notes