

Emma Oxford

## Survey data

### When what you see is not what you wanted to get

**S**urveys are an easy way to collect a lot of information fast, and libraries are always in need of information from their patrons, their staff, community partners, and more. Given the prevalence of survey-based research in the literature (a search of just the word *survey* on LISTA yields almost 28,000 results), it is time the library community took a more in-depth look at this research tool.

#### Common problems with survey data

Most problems with surveys stem from the fact that no survey respondent will ever care about the accuracy of the results as much as the researcher does. (Even non-response bias comes from this.) We've all been in this position: you get an email on a listserv, a request to fill out a "ten-minute survey" that you will breeze through in five, so you fill out the multiple-choice questions, and you skip the freeform ones, and two minutes later you pat yourself on the back for doing your good deed for the day, because, hey, you gave the researchers some data, and surely some data is better than no data, right?

Wrong. Some data is only better than no data if it is good data. Bad data is never better than no data. We all know this because we've all been on the other side of the survey, sending out well-intentioned questionnaires hoping for feedback so that we can make well-informed decisions and then being dismayed when half-hearted responses come back to us.

Filling out a survey haphazardly, looking for shortcuts, or not giving real thought

to the questions is known as *satisficing*, a term borrowed from economics.<sup>1</sup> Satisficing can take the form of many different behaviors, including nondifferentiation (choosing the same response to every question), skipping items, rushing, or quitting early.<sup>2</sup> Research in other disciplines has examined the prevalence of satisficing and its impact on survey data. Scott Barge and Hunter Gehlbach report that in two surveys of university students, the majority of respondents engaged in some form of satisficing.<sup>3</sup> It is important to verify that these responses have not had an adverse effect on the data analysis before making any data-driven decisions.

#### My own experience

Satisficing happens in library research, but based on my own experience, there is another problem at hand: real misinformation being reported. In the spring of 2015, I sent a survey on reference services to librarians at a group of North American liberal arts institutions. The survey was inherently complex (although I believed the questions quite simple) because it collected both institutional data (e.g., "Does your institution have more than one library?") and individual data (e.g., "How satisfied are you with reference services at your institution?"). In order to capture the individual data, I allowed for responses from

---

Emma Oxford is science librarian at Rollins College's Olin Library, email: [eoxford@rollins.edu](mailto:eoxford@rollins.edu)

© 2016 Emma Oxford

---

**Some data is only better than no data if it is good data. Bad data is never better than no data.**

---

multiple librarians per institution. I received 121 responses from 60 distinct institutions.

While my method yielded useful results in both respects, it also presented some problems at the institutional level. Specifically, in some instances, respondents from the same institution could not agree on basic factual information.

In two instances, they disagreed on whether their institution had more than one library. Even when all the respondents agreed on how many libraries they had, in six instances they then disagreed on which of these libraries had designated reference desks (a term defined at the beginning of the survey); eleven times they disagreed on how many desks there were in total; twelve times they disagreed on how many hours the desk(s) were staffed, and in fourteen instances they disagreed on who was staffing the desk(s) (students, staff, librarians, or various combinations of the three). This isn't even getting into questions on the other types of reference services their library offered.

Yikes.

To be fair, in some of these instances respondents provided comments that clarified their answers and resolved the discrepancies. Additionally, such disagreement only occurred in a minority of responses. But it is a significant minority. You might say this is what I get for trying to collect institutional data while allowing for multiple respondents per institution. But in the instances where only one librarian responded, how do I know they wouldn't have been at odds with their colleagues had other librarians from that institution weighed in?

Now, as the first study I ever conducted as a professional librarian, this survey was by no means the best designed survey in the history of library research. Some of the fault

for the conflicting results lies with me and the survey design. And some relevant data did come out of the survey, but that doesn't mean the implications of the discrepancies can be ignored.

### **Considerations for the library community**

So what are the implications? Is this a simple case of satisficing? Or is this a real case of librarians at the same institution not being on the same page? If libraries are going to use surveys as a dominant mode of data collection, they need to be relatively certain that the results of any given survey are accurate and meaningful.

There is no such thing as a perfect survey, and there is no such thing as a flawless dataset. There will always be problems, and we know some common ones to expect. But I wouldn't have expected academic librarians to disagree on how many libraries their campus had, and this only came to light because I permitted multiple responses from the same institution. In most cases, when a researcher is collecting institutional data, they will only have one response per institution. How do we weed out bad data when it looks just like the good data?

Survey takers can also take something away from this. Most importantly, having no data is better than having bad data. If you are not willing to give your full attention to a survey and to answer the questions to the best of your ability, don't take it. You will actually be making the life of the researcher a lot easier.

### **Notes**

1. Yavuz Akbulut, "Predictors of Inconsistent Responding in Web Surveys," *Internet Research* 25, no. 1 (2015): 131-47, doi:10.1101/IntR-01-2014-0017.
2. Scott Barge and Hunter Gehlbach, "Using the Theory of Satisficing to Evaluate the Quality of Survey Data," *Research in Higher Education* 53, no. 2 (2012): 182-200, doi:10.1007/s11162-011-9251-2.
3. Ibid. *zz*