

 Lagunita is retiring and will shut down at 12 noon Pacific Time on March 31, 2020. A few courses may be open for self-enrollment for a limited time. We will continue to offer courses on other online learning platforms; visit <http://online.stanford.edu>.

Course > Producing Data: Designing Studies > Observational Studies > Sample Surveys: Design Issues

 Bookmark this page

## Sample Surveys: Design Issues

**Learning Objective: Determine how the features of a survey impact the collected data and the accuracy of the data.**

### Sample Surveys

A sample survey is a particular type of observational study in which individuals report variables' values themselves, frequently by giving their opinions. Researchers have several options to choose from when deciding how to survey the individuals involved: in person, or via telephone, Internet, or mail.

The following issues in the design of sample surveys will be discussed:

- open vs. closed questions
- unbalanced response options
- leading questions
- planting ideas with questions
- complicated questions
- sensitive questions

These issues are best illustrated with a variety of concrete examples.

Suppose you want to determine the musical preferences of all students at your university, based on a sample of students. In the **Sampling** section, we discussed various ways to obtain the sample, such as taking a simple random sample from all students at the university, then contacting the chosen subjects via email to request their responses and following up with a second email to those who did not

respond the first time. This method would ensure a sample that is fairly representative of the entire population of students at the university, and avoids the bias that might result from a flawed designs such as a convenience sample or a volunteer sample.

However, even if we managed to select a representative sample for a survey, we are not yet home free: we must still compose the survey question itself so that the information we gather from the sampled students correctly represents what is true about their musical preferences. Let us consider some possibilities:

**Question:** "What is your favorite kind of music?"

This is what we call an **open question**, which allows for almost unlimited responses. It may be difficult to make sense of all the possible categories and subcategories of music that survey respondents could come up with. Some may be more general than what you had in mind ("I like modern music the best") and others too specific ("I like Japanese alternative electronic rock by Cornelius"). Responses are much easier to handle if they come from a **closed question**:

**Question:** Which of these types of music do you prefer: classical, rock, pop, or hip-hop?

What will happen if a respondent is asked the question as worded above, and he or she actually prefers jazz or folk music or gospel? He or she may pick a second-favorite from the options presented, or try to pencil in the real preference, or may just not respond at all. Whatever the outcome, it is likely that overall, the responses to the question posed in this way will not give us very accurate information about general music preferences. If a closed question is used, then great care should be taken to include all the reasonable options that are possible, including "not sure." Also, in case an option was overlooked, "other: \_\_\_\_\_" should be included for the sake of thoroughness.

Many surveys ask respondents to assign a rating to a variable, such as in the following:

**Question:** How do you feel about classical music? Circle one of these: I love it, I like it very much, I like it, I don't like it, I hate it.

Notice that the options provided are rather "top-heavy," with three favorable options vs. two unfavorable. If someone feels somewhat neutral, they may opt for the middle choice, "I like it," and a summary of the survey's results would distort the respondents' true opinions.

Some survey questions are either deliberately or unintentionally biased towards certain responses:

**Question:** "Do you agree that classical music is the best type of music, because it has survived for centuries and is not only enjoyable, but also intellectually rewarding? (Answer yes or no.)"

This sort of wording puts ideas in people's heads, urging them to report a particular opinion. One way to test for bias in a survey question is to ask yourself, "Just from reading the question, would a respondent have a good idea of what response the surveyor is hoping to elicit?" If the answer is yes, then the question should have been worded more neutrally.

Sometimes, survey questions are ordered in such a way as to deliberately bias the responses by planting an idea in an earlier question that will sway people's thoughts in a later question.

**Question:** In the year 2002, there was much controversy over the fact that the Augusta National Golf Club, which hosts the Masters Golf Tournament each year, does not accept women as members. Defenders of the club created a survey that included the following statements. Respondents were supposed to indicate whether they agreed or disagreed with each statement:

"The First Amendment of the U.S. Constitution applies to everyone regardless of gender, race, religion, age, profession, or point of view."

"The First Amendment protects the right of individuals to create a private organization consisting of a specific group of people based on age, gender, race, ethnicity, or interest."

"The First Amendment protects the right of organizations like the Boy Scouts, the Girls Scouts, and the National Association for the Advancement of Colored People to exist."

"Individuals have a right to join a private group, club, or organization that consists of people who share the same interests and personal backgrounds as they do if they so desire."

"Private organizations that are not funded by the government should be allowed to decide who becomes a member and who does not become a member on their own, without being forced to take input from other outside people or organizations."

Notice how the first and second statements steer people to favor the opinion that specialized groups may form private clubs. The third statement reminds people of organizations that are formed by groups on the basis of gender and race, setting the stage for them to agree with the fourth statement, which supports people's rights to join any private club. This in turn leads into the fifth statement, which focuses on a private organization's right to decide on its membership. As a group, the questions attempt to relentlessly steer a respondent towards ultimately agreeing with the club's right to exclude women.

Sometimes surveyors attempt to get feedback on more than one issue at a time.

**Question:** "Do you agree or disagree with this statement: 'I don't go out of my way to listen to modern music unless there are elements of jazz, or else lyrics that are clear and make sense.'"

Put yourself in the place of people who enjoy jazz and straightforward lyrics, but don't have an issue with music being "too modern," per se. The logic of the question (or lack thereof) may escape the respondents, and they would be too confused to supply an answer that correctly conveys their opinion. Clearly, simple questions are much better than complicated ones; rather than try to gauge opinions on several issues at once, complex survey questions like this should be broken down into shorter, more concise ones.

Depending on the topic, we cannot always assume that survey respondents will answer honestly.

**Question1:** "Have you eaten rutabagas in the past year?"

If respondents answer no, then we have good reason to believe that they did not eat rutabagas in the past year.

**Question2:** "Have you used illegal drugs in the past year?"

If respondents answer no, then it is still a possibility that they did use illegal drugs, but didn't want to admit it.

Effective techniques for collecting accurate data on sensitive questions are a main area of inquiry in statistics. One simple method is **randomized response**, which allows individuals in the sample to answer anonymously, while the researcher still gains information about the population. This technique is best illustrated by an example.

**Example**

For the question, "Have you used illegal drugs in the past year?" respondents are told to flip a fair coin (in private) before answering and then answer based on the result of the coin flip: if the coin flip results in "Heads," they should answer "Yes" (regardless of the truth), if a coin flip results in "Tails," they should answer truthfully. Thus, roughly half of the respondents are "truth-tellers," and the other half give the uncomfortable answer "Yes," without the interviewer's knowledge of who is in which group. The respondent who flips "Tails" and answers truthfully knows that he or she cannot be distinguished from someone who got "Heads" in the coin toss. Hopefully, this is enough to encourage respondents to answer truthfully. As we will learn later in the course, the surveyor can then use probability methods to estimate the proportion of respondents who admit they used illegal drugs in this scenario, while being unable to identify exactly which respondents have been drug abusers.

Besides using the randomized response method, surveyors may encourage honest answers from respondents in various other ways. Tactful wording of questions can be very helpful. Giving people a feeling of anonymity by having them complete questionnaires via computer, rather than paper and pencil, is another commonly used technique.

Open Learning Initiative [↗](#)



[↗](#) Unless otherwise noted this work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License [↗](#).

© All Rights Reserved