

Nonresponse and Response Bias

by Sophia Tutorial



WHAT'S COVERED

This tutorial will cover the topics of nonresponse bias and response bias by focusing on:

1. Nonresponse Bias
2. Participation Bias
3. Response Bias

1. Nonresponse Bias

A nice way to think of sampling is to use a "pot of soup" analogy. You want a representative sample, right? Well, you don't need to drink the entire pot of soup in order to figure out what's in it. You just need the right taste.

It would be like selecting all of the ingredients from the soup in a single tasting, but certain things can go wrong with the taste test that can affect what you think is in the soup. Just like you don't really know what the population looks like, you really don't have a clear idea of all the ingredients in the soup. All you get is the taste, and if you don't get the right taste, you're going to leave something out and not know exactly what's in the soup (or, population).

In terms of sampling, **nonresponse** means that someone selected for the sample either can't be contacted or is unwilling to participate.

Now, nonresponse happens. It's an inevitability that you will get uncooperative people, people that don't want to take your survey or people who refuse to be part of your experiment. It may be that you just won't be able to contact certain people.

The problem of nonresponse is not a problem until the people that weren't able to be contacted or refused to participate differ substantially from the people that were in the sample. Now the sample is not representative of the population. That is called **nonresponse bias** because you're not getting an accurate cross-section of opinions. The opinions of people that you wanted to get are left out.

IN CONTEXT

A workplace wishes to survey 200 of its 1,000 employees about their workload and their stress level,

so they put 200 surveys in the workers' mailboxes. It's likely that the people who have the biggest workloads might get left out of the sample because they don't check their mailboxes as often as other people. Or if they do get around to checking their mailbox, they may not complete the survey, or don't return it, because they're so busy.

What effect might that have? The 200 respondents that completed the survey may have reported that workload level is not that high. The only problem is that the people with the lower workloads are the only people who turned them in, because they had the time to take it. Also, the people with the higher workloads didn't have the time to take it, reinforcing the conclusion that the company might think the workload level is lower than it really is.

The nonresponse rate is easy to calculate. You just subtract the number that you got back from the number that you mailed out, and that's your nonresponse rate.

➦ **EXAMPLE** Say you mailed out 100, and you only got 80 back. Well, that's 20 out of 100, or 20% nonresponse rate.



THINK ABOUT IT

Consider the different ways of conducting a survey, a poll, or a sample. Which of the following methods do you think has the highest nonresponse rate?

- Mail
- Telephone
- Face-to-Face

The answer is the mail. People will either throw it away, forget to fill it out, or maybe they'll fill it out and then forget to mail it back. This is problematic because when the United States takes its census of everyone in the country, it does so by mail. Sometimes they have to do follow-ups.

In samples with high rates of nonresponse, follow-ups typically are needed. Suppose you started with a mailing. You might need to follow up by calling them at home. If you can't reach them by calling them at home, you might need to follow up by coming directly to their house.

Sometimes, even when they are contacted, someone will refuse to participate. Follow-ups like this might be more necessary in some areas of the country than others because different areas of the country have different rates of nonresponse.



TERMS TO KNOW

Nonresponse

Nonresponse is a lack of response from people you've selected. It affects the ability to draw conclusions from your sample.

Nonresponse Bias

Bias that occurs when the people who were unable to be reached or unwilling to participate in a sample have substantially different opinions than the people who were included in the sample, resulting in a misrepresentation of the population.

2. Participation Bias

On the other end of the spectrum is when people are excessively passionate about a topic and they're eager to participate. The people who raise their hand to participate are volunteering their time because they have a strong opinion about the topic at hand. **Participation bias** happens when people participate because they have strong opinions about the topic, or they're ambivalent because they are only participating because they are getting paid to participate.

🔗 **EXAMPLE** Suppose you need to gather information on an upcoming election and you ask people to participate in a focus group. In your group, you find that you have a group in strong support of the Democratic party and you have a group in strong support of the Republican party, and no one in the middle.

To correct this, you decide you're going to pay participants \$20 for their time. Now your group is filled with people who will simply tell you what they think you want to hear, which invites participation bias.



TERM TO KNOW

Participation Bias

Bias that occurs when participation in a study is voluntary. People who feel strongly may be the only participants.

3. Response Bias

Response bias is when people's answers are influenced. Remember the pot of soup analogy? When you get a representative sample, that's like getting a little taste of everything in the soup. However, things can go wrong and you don't get the right taste of the soup.

Response bias can occur if the wording of the question is unclear to the respondent, if a respondent is uncomfortable due to the sensitive or personal nature of the questions, or if the respondent feels like the questioner is implying that the question has a "correct" response. That's also called social desirability bias.

IN CONTEXT

On April 20, 1993, the *New York Times* published an article on a survey conducted by the Roper Organization on behalf of the Jewish American Community about the soon-to-be opened Holocaust Museum in Washington, DC.

The newspaper reported that 22%, an astounding number of adults surveyed, expressed some doubt as to whether the Holocaust had actually occurred. The actual question that was presented to people was:

"Does it seem possible, or does it seem impossible to you, that the Nazi extermination of the Jews never happened?"

This seems to be a fairly straightforward question, but there was a big problem with it, and it caused response bias. The problem is that the question contained a double negative, which are confusing. Saying it is *impossible* that it *never* happened is the same as someone saying they are certain that it *did* happen, but the question doesn't clearly read that way.

The good thing is that, one year later, the question was revised, and it became clearer. The new question stated:

"Does it seem possible to you that the Nazi extermination of the Jews never happened, or do you feel certain that it happened?"

With this new, clearer question, the question clearly distinguishes between what the two options are--"does it seem possible," or "do you feel certain?" With the two options clearly defined, less than 2% of individuals were unsure as to whether it was real or not. This provided a more accurate interpretation of what the American public felt.

Therefore, unclear questions can lead to an inaccurate representation due to response bias. The other scenario in which this can occur is when people will answer a question because they are either ashamed, or they think that there's a "right" answer that someone is fishing for.

There are certain topics that are particularly sensitive and might make a person want to lie.

Topics that Could Result in a Response Bias	
Drugs	This may result in many people saying they've never used drugs, whether they actually have or not. Even if there's no consequence and the survey is anonymous, they'll still say they've never used drugs when, in fact, they have.
Criminal history	Participants might say they don't have one, even if they do.
Sexual behavior	This might cover topics of a highly sensitive and personal nature.
Racial prejudice	There's an implied right answer; people don't want to say that they're racially prejudiced.
Income	People will report it as being higher than it actually is if they're of low-income status, or even possibly more surprisingly, people will report it as lower than it really is if they're of very high-income status. A lot of people don't want to be showy about their wealth, and so they'll try and come up with a more reasonable number, in their eyes.

How does this affect what we think about the population? How does this affect the "soup?"

It's like taking a sample of the soup and only tasting the things that you want to taste. Maybe you don't like beans, and so you just sort of ignore the fact that they're in there. You don't get the overall flavor of what's supposed to happen. It's the same thing with response bias. It doesn't give you the right overall interpretation of what things the population is supposed to be like.



TERM TO KNOW

Response Bias

Bias that occurs when either (1) the question is poorly worded so that certain responses are over-represented, or (2) the respondent is confused by the question or feel like they should lie due to the sensitive nature of the question.



SUMMARY

Nonresponse bias occurs when people who are selected for the sample can't participate, either because you can't find them, or because they're actively refusing. The biggest problem is that if you have high rates of nonresponse, it might give you an inaccurate representation of what's going on with your population. You won't be able to use your sample to draw an inference about your population. Response bias occurs one of two ways: either a respondent doesn't understand the question and so gives an answer that he wasn't intending; or, the respondent wants to give a supposedly correct answer to the questioner. Both of these can be inaccurate representations of what actually is the truth about the population. Response bias is a tough thing to get rid of, especially when it is unintentional and surrounds the wording of the questions.

Good luck!

Source: Adapted from Sophia tutorial by Jonathan Osters.



TERMS TO KNOW

Nonresponse Bias

Bias that occurs when the people who were unable to be reached or unwilling to participate in a sample have substantially different opinions than the people who were included in the sample, resulting in a misrepresentation of the population.

Participation Bias

Bias that occurs when participation in a study is voluntary. People who feel strongly may be the only participants.

Response Bias

Bias that occurs when either (1) the question is poorly worded so that certain responses are over-represented, or (2) the respondent is confused by the question or feel like they should lie due to the sensitive nature of the question.