

## 20.4: Conducting surveys

under construction, missing citations

### Introduction

What is a survey? A **survey** is a method of collecting information from a sample of a **reference population**. Surveys are implemented in many fields including biomedical work (see [Chapter 5.4](#)). We can design a **cross-sectional** study, which gathers information at one time, or the study can be **longitudinal**, whereby information is gathered over a period of time. If the purpose of the survey is to determine **association** between two (or more) variables, then either cross-sectional or longitudinal approaches will do. However, if **cause-effect hypotheses** are the purpose, then longitudinal (e.g., **prospective cohort**) approaches would be better.

### Survey basics

Steps to conduct a survey include

1. Identify and clarify the purpose of the survey. If the purpose is to find out how common something is, then this is descriptive. If we are interested in why something has occurred, then this is an analytic survey.
2. Define the reference population. It is essential that you know which group the survey applies to. For example, if one wishes to study the opinions of undergraduates at your university, then postgraduate students cannot be included in the sample as they are not part of the reference population.
3. Design sampling method and determine sample size. Sampling needs to be done to obtain unbiased sample from representative population. If the size of the population is known, then a target of 10% might be the relevant sample size, and a procedure should be taken to obtain a simple random sample. A measure of the success of a survey sample design is the size of the response rate, defined as the ratio of surveys returned divided by the number of surveys distributed.
4. What information is needed? Care needs to be taken to make sure that the questions asked actually yield the desired information. For example, if the questionnaire is long, there may be a tendency for some to skim or skip questions. If too much information is requested, this may lower the response rate.
5. How will the information be collected? Types and format of questions (closed or open ended). A phone survey? Written survey dropped off as a mailer? Interview?
6. In thinking about the data to be collected from the questions, you also need some scoring system. Will a dichotomous response (e.g., True/False, or Yes/No) be adequate, or would a Likert-like scale be more appropriate? When scaling responses, one needs to also be concerned with floor and ceiling effects.
7. Collect the data. What protocol will be employed to achieve a high response rate with unbiased responses?
8. Analyze the data. Often chi-square contingency table or the related logistic regression would be appropriate.

### Bias sources in survey research

When a statistical estimator consistently under or over estimates the true value, this is called **statistical bias**, which was introduced and discussed in [Chapter 5.3](#). The potential for bias responses to survey questions is an important constraint on the applicability of the survey. An example, it is well-known that adults tend to overestimate their height, but underestimate their weight.

In survey research, different classes of bias have been defined.

- **Information bias** occurs when trends are present in the measurement of the response, (1) recall bias and (2) observer bias.
- **Recall bias** is when a difference occurs because some people are much more likely to remember and event than others.
- **Observer bias** can be as a result of differences between different observers. If different persons conduct interviews, then it is important that all observers use a standardized method of collecting data.
- A reality of survey is that not all targets of the survey will answer the questions. This may result in bias. **Non-response bias** is the situation when those who respond to a questionnaire, the **responders**, differ in some way from those who don't, the **non-responders**.
- **Selection bias** results when the sample group you have chosen is not representative of the population you want to generalize your results to. Random sampling can help to minimize this from happening in your survey, but a stratified sampling approach is needed to avoid missing representation, e.g., economic groups, ethnicity.

### How to ask questions?

The goal is to maximize the number of people who respond to the survey while maintaining accuracy and relevance of the responses. This is accomplished by asking the right questions in the right manner, but also by how the questionnaire is presented

and administered. To get accurate answers, one should include additional questions to check the consistency of the responses provided by a person. For example, if the study is about smoking, you could ask either of two questions (or both...?):

**Question 1.** Do you smoke tobacco cigarettes? Yes / No

**Question 2.** How many cigarettes did you smoke yesterday?

0 .... 1 – 10 .... 11 – 20 .... 21 +

Note that these questions are CLOSED — the responder must answer using the answers provided rather than making something up. This has the advantage of restricting the possible answers and allows you to test specific hypotheses. An OPEN question might be

**Question 3.** Do you smoke a lot of cigarettes in a day?

As you can imagine, you would expect to get a variety of interpretations of this question, which limits your ability to analyze test the hypothesis. It would be a poor question to use.

Closed or forced format questions can take on a variety of styles.

**Question 4.** What is your favorite soft drink? (select one answer only)

- ☐ water
- ☐ cola
- ☐ ice tea
- ☐ fruit juice
- ☐ no preference

Note that Question 4 gives the responder a choice among categories. For another example,

**Question 5.** Biostatistics is my favorite subject

☐ Strongly disagree ☐ Disagree ☐ Undecided ☐ Agree ☐ Strongly Agree

And still another example might employ ranking, requesting the responder to rank from 1 to 8 their favorite subjects from a list of academic topics.

All of these closed format responses can be easily converted for analyses by contingency table or other nonparametric statistics.

### Some other general tips in writing a survey.

- Keep sentences simple and short.
- Ask for only one piece of information at a time.
- Ask precise questions.
- Start the survey with the question(s) most relevant to the subject of the study.
- Avoid asking personal questions at the start.
- Write some questions then conduct a pilot study to test for the efficacy of the survey

### Suggested readings

[Statistics Canada](#): a site with lots of material about survey methods.

[Wikipedia: Statistical survey](#)

---

This page titled [20.4: Conducting surveys](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Michael R Dohm](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.