

## Section 9.2: Potential Problems with collecting Data

There are several factors that might lead to problems with data collection.

Potential Problem	What It Means	Example
Bias	The question influences responses in favour of, or against the topic of the data collection.	Suppose a person asks: <i>Don't you think the price of a movie ticket is too high?</i> This person has a bias against the current ticket price, and the bias influences how the survey question is written.
Use of Language	The use of language in a question could lead people to give a particular answer.	If you ask: <i>Don't you think the price of a movie ticket is too high?</i> , the question may lead people to answer yes. A better question would be: <i>Do you think the price of a movie ticket is too high, too low, or fair?</i>
Timing	<i>When</i> the data are collected could lead to particular results.	A survey is conducted to find opinions on the need for a vehicle to have winter tires. The results may be different if the survey is conducted in August instead of February.
Privacy	If the topic of the data collection is personal, a person may not want to participate or may give an untrue answer on purpose. Anonymous surveys may help.	People may not want to participate in a study on weight if it means stepping on a scale in front of other people.

Cultural Sensitivity	Cultural sensitivity means that you are aware of other cultures. You must avoid being offensive and asking questions that do not apply to that culture.	Suppose you wanted to know the favourite method of cooking ham, and you asked: <i>Please circle your favourite method:</i> <i>BBQ                      Bake                      Fry</i> This question does not apply to everyone because many people do not eat ham. A better question would be: <i>If you eat ham, name your favourite method of cooking it.</i>
Ethics	Ethics dictate that collected data must not be used for purposes other than those told to the participants. Otherwise, your actions are considered unethical.	Suppose you tell your classmates that you want to know their favourite snacks to help you plan your birthday party. If you then use the information to try to sell your classmates their favourite snacks between classes, your actions would be unethical.
Cost	The cost of collecting data must be taken into account.	If you need to pay for printing the questionnaires, or to pay people to collect the data, the cost may be more than you can afford.
Time	The time needed for collecting the data must be considered.	A survey that takes an hour to complete may be too long for most people. This would limit the number of people willing to participate.

**Example 1** Identifying and Eliminating Potential Problems

For each survey question, explain why a problem may occur and the effect it would have on the results. Suggest how each problem could be avoided.

- a) A survey is conducted to find out if citizens think the local government should provide more money for youth activities. The question asked was: “Would you support an increase in taxes to create more skate parks?”
- b) A survey is conducted to find out the level of school spirit. Students are polled about their level of school spirit after the soccer team wins the championship.

**A Solution**

- a) The use of language in the survey question could be problematic.  
The question emphasizes what citizens would lose; that is, their taxes would increase. The question also downplays what citizens would gain by only mentioning skate parks, instead of a variety of activities.  
Most people would probably respond by saying they would not support an increase in taxes to build more skate parks.  
A better question would be: “Do you think the local government should supply more funds for youth recreational activities?”
- b) The timing of the survey question could be problematic.  
Since the school’s soccer team just won the championship, the level of school spirit would be higher than usual.  
The results of the survey may show a higher level of school spirit than if the survey was conducted at another time.  
Asking students the same question a month later, when no school event is occurring, should produce more accurate results.

**Example 2** Analyzing Data Collection for Problems

Kublu and Irniq plan to open a shop in Saskatoon that would sell traditional Inuit crafts.

To ensure Saskatoon is the best place for their business, they want to survey residents to find out how popular Inuit crafts are.

Kublu knows that they would get the most accurate results if each household in Saskatoon is surveyed, but Irniq points out that this is problematic.

Explain why.

**▶ A Solution**

The number of households in Saskatoon is great. Kublu and Irniq may have problems related to cost and time.

The cost of printing and mailing enough surveys for each household would be very high. Also, Kublu and Irniq should provide an envelope and stamp for each household to return the survey. This would be an additional cost.

The time it would take to print, mail, and collect the surveys for all the households would be too long.

**Example 3****Overcoming Potential Problems of Data Collection**

Antonia wants to find out if there is a relationship between household income and how much people spent on Christmas presents.

Identify potential problems Antonia may encounter, and explain how she could deal with the problems.

**A Solution**

Christmas is not celebrated by all cultures, and so the survey question does not apply to everyone. An appropriate opening question for the survey might be: "Do you celebrate Christmas?" If a person responds "No", then he or she will not need to answer the other question in your survey.

Information about income and spending habits is personal, so people may be uncomfortable revealing it. An anonymous survey would be appropriate.

The use of language may influence responses.

Examples of inappropriate or intrusive questions would be:

"How much do you make?" and "How much do you spend?"

A better question might be:

"Is the amount you spend on Christmas presents:

- greater than your weekly income?
- less than your weekly income?
- equal to your weekly income?"

Questions: p435  
#3, 6, 8



3. Name a problem with each data collection.
- After the first week of school, your principal asks you and your friends how you are enjoying school.
  - An online magazine asks readers either to agree or disagree with the statement: "If you find a \$20-bill, you turn it in."
  - Brenda asks her classmates if they think girls should not be allowed to cover their heads in school.
  - To discover the most popular kind of movie at his school, Carlos plans to ask each student what her or his favourite kind of movie is.

(a) Privacy

(b) Use of Language

(c) Cultural Sensitivity

(d) Time

6. Andrew went to each class in his school and asked for a show of hands to find out how many students had ever been bullied at school. Only 2 students raised their hands.

Andrew concluded that bullying was not a major problem at his school.

- Is this a reasonable conclusion? Explain.
  - Describe a better method for conducting Andrew's survey.
8. a) Describe how each question reveals a bias of the questioner.
- Do you think it is a good idea to use DNA tests to convict a violent criminal?
  - Do you think gas guzzling SUVs should be banned?
  - Do you think students should be allowed to use spell check because it automatically improves spelling?

(a) No, many students who are bullied are afraid to tell people, especially in front of their classmates.

(b) An anonymous survey

(i) The use of the words 'violent criminals' is bias towards using DNA tests.

(ii) The use of 'gas guzzling' is negative towards SUV's.

(iii) The question emphasizes the positive aspect of spell checks.

