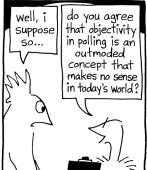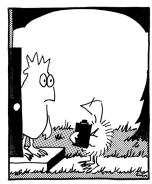## 2.4 Bias in Surveys

The results of a survey can be accurate only if the sample is representative of the population and the measurements are objective. The methods used for choosing the sample and collecting the data must be free from **bias**. Statistical bias is any factor that favours certain outcomes or responses and hence systematically skews the survey results. Such bias is often unintentional. A researcher may inadvertently use an unsuitable method or simply fail to recognize a factor that prevents a sample from being fully random. Regrettably, some people deliberately bias surveys in order to get the results they want. For this reason, it is important to understand not only how to use statistics, but also how to recognize the misuse of statistics.

INVESTIGATE & INQUIRE: Bias in a Survey

1. What sampling technique is the pollster in this cartoon likely to be using?
2. What is wrong with his survey methods? How could he improve them?
3. Do you think the bias in this survey is intentional? Why or why not?
4. Will this bias seriously distort the results of the survey? Explain your reasoning.
5. What point is the cartoonist making about survey methods?
6. Sketch your own cartoon or short comic strip about data management.

Sampling bias occurs when the sampling frame does not reflect the characteristics of the population. Biased samples can result from problems with either the sampling technique or the data-collection method.

Example 1 Sampling Bias

Identify the bias in each of the following surveys and suggest how it could be avoided.

a) A survey asked students at a high-school football game whether a fund for extra-curricular activities should be used to buy new equipment for the football team or instruments for the school band.

b) An aid agency in a developing country wants to know what proportion of households have at least one personal computer. One of the agency's staff members conducts a survey by calling households randomly selected from the telephone directory.

Solution

a) Since the sample includes only football fans, it is not representative of the whole student body. A poor choice of sampling technique makes the results of the survey invalid. A random sample selected from the entire student body would give unbiased results.

b) There could be a significant number of households without telephones. Such households are unlikely to have computers. Since the telephone survey excludes these households, it will overestimate the proportion of households that have computers. By using a telephone survey as the data-collection method, the researcher has inadvertently biased the sample. Visiting randomly selected households would give a more accurate estimate of the proportion that have computers. However, this method of data collection would be more time-consuming and more costly than a telephone survey.

Non-response bias occurs when particular groups are under-represented in a survey because they choose not to participate. Thus, non-response bias is a form of sampling bias.

Example 2 Non-Response Bias

A science class asks every fifth student entering the cafeteria to answer a survey on environmental issues. Less than half agree to complete the questionnaire. The completed questionnaires show that a high proportion of the respondents are concerned about the environment and well-informed about environmental issues. What bias could affect these results?

Solution

The students who chose not to participate in the survey are likely to be those least interested in environmental issues. As a result, the sample may not be representative of all the students at the school.

To avoid non-response bias, researchers must ensure that the sampling process is truly random. For example, they could include questions that identify members of particular groups to verify that they are properly represented in the sample.

Measurement bias occurs when the data-collection method consistently either under- or overestimates a characteristic of the population. While random errors tend to cancel out, a consistent measurement error will skew the results of a survey. Often, measurement bias results from a data-collection process that affects the variable it is measuring.

Example 3 Measurement Bias

Identify the bias in each of the following surveys and suggest how it could be avoided.

a) A highway engineer suggests that an economical way to survey traffic speeds on an expressway would be to have the police officers who patrol the highway record the speed of the traffic around them every half hour.

b) As part of a survey of the "Greatest Hits of All Time," a radio station asks its listeners: Which was the best song by the Beatles?

    i) Help!   **ii)** Nowhere Man

    iii) In My Life   **iv)** Other:

c) A poll by a tabloid newspaper includes the question: "Do you favour the proposed bylaw in which the government will dictate whether you have the right to smoke in a restaurant?"

Solution

a) Most drivers who are speeding will slow down when they see a police cruiser. A survey by police cruisers would underestimate the average traffic speed. Here, the data-collection method would systematically decrease the variable it is measuring. A survey by unmarked cars or hidden speed sensors would give more accurate results.

b) The question was intended to remind listeners of some of the Beatles' early recordings that might have been overshadowed by their later hits. However, some people will choose one of the suggested songs as their answer even though they would not have thought of these songs without prompting. Such **leading questions** usually produce biased results. The survey would more accurately determine listeners' opinions if the question did not include any suggested answers.

c) This question distracts attention from the real issue, namely smoking in restaurants, by suggesting that the government will infringe on the respondents' rights. Such **loaded questions** contain wording or information intended to influence the respondents' answers. A question with straightforward neutral language will produce more accurate data. For example, the question could read simply: "Should smoking in restaurants be banned?"

Response bias occurs when participants in a survey deliberately give false or misleading answers. The respondents might want to influence the results unduly, or they may simply be afraid or embarrassed to answer sensitive questions honestly.

Example 4 Response Bias

A teacher has just explained a particularly difficult concept to her class and wants to check that all the students have grasped this concept. She realizes that if she asks those who did not understand to put up their hands, these students may be too embarrassed to admit that they could not follow the lesson. How could the teacher eliminate this response bias?

Solution

The teacher could say: "This material is very difficult. Does anyone want me to go over it again?" This question is much less embarrassing for students to answer honestly, since it suggests that it is normal to have difficulty with the material. Better still, she could conduct a survey that lets the students answer anonymously. The teacher could ask the students to rate their understanding on a scale of 1 to 5 and mark the ratings on slips of paper, which they would deposit in a box. The teacher can then use these ballots to decide whether to review the challenging material at the next class.

As the last two examples illustrate, careful wording of survey questions is essential for avoiding bias. Researchers can also use techniques such as follow-up questions and guarantees of anonymity to eliminate response bias. For a study to be valid, all aspects of the sampling process must be free from bias.

Key Concepts

- Sampling, measurement, response, and non-response bias can all invalidate the results of a survey.

- Intentional bias can be used to manipulate statistics in favour of a certain point of view.

- Unintentional bias can be introduced if the sampling and data-collection methods are not chosen carefully.

- Leading and loaded questions contain language that can influence the respondents' answers.

1. Explain the difference between a measurement bias and a sampling bias.

2. Explain how a researcher could inadvertently bias a study.

3. Describe how each of the following might use intentional bias
   a) the media
   b) a marketing department
   c) a lobby group

Practise

A

1. Classify the bias in each of the following scenarios.

   a) Members of a golf and country club are polled regarding the construction of a highway interchange on part of their golf course.

   b) A group of city councillors are asked whether they have ever taken part in an illegal protest.

   c) A random poll asks the following question: "The proposed casino will produce a number of jobs and economic activity in and around your city, and it will also generate revenue for the provincial government. Are you in favour of this forward-thinking initiative?"

   d) A survey uses a cluster sample of Toronto residents to determine public opinion on whether the provincial government should increase funding for the public transit.

Apply, Solve, Communicate

2. For each scenario in question 1, suggest how the survey process could be changed to eliminate bias.

3. **Communication** Reword each of the following questions to eliminate the measurement bias.

   a) In light of the current government's weak policies, do you think that it is time for a refreshing change at the next federal election?

   b) Do you plan to support the current government at the next federal election, in order that they can continue to implement their effective policies?

   c) Is first-year calculus as brutal as they say?

   d) Which of the following is your favourite male movie star?

   i) Al Pacino   **ii)** Keanu Reeves

   iii) Robert DeNiro   **iv)** Jack Nicholson

   v) Antonio Banderas **vi)** Other:

   e) Do you think that fighting should be eliminated from professional hockey so that skilled players can restore the high standards of the game?

B

4. **Communication**

   a) Write your own example of a leading question and a loaded question.

   b) Write an unbiased version for each of these two questions.

5. A school principal wants to survey data-management students to determine whether having computer Internet access at home improves their success in this course.

a) What type of sample would you suggest? Why? Describe a technique for choosing the sample.

b) The following questions were drafted for the survey questionnaire. Identify any bias in the questions and suggest a rewording to eliminate the bias.

i) Can your family afford high-speed Internet access?

ii) Answer the question that follows your mark in data management. Over 80%: How many hours per week do you spend on the Internet at home?
60–80%: Would home Internet access improve your mark in data management ?
Below 60%: Would increased Internet access at school improve your mark in data management?

c) Suppose the goal is to convince the school board that every data-management student needs daily access to computers and the Internet in the classroom. How might you alter your sampling technique to help achieve the desired results in this survey? Would these results still be statistically valid?

6. Application A talk-show host conducts an on-air survey about re-instituting capital punishment in Canada. Six out of ten callers voice their support for capital punishment. The next day, the host claims that 60% of Canadians are in favour of capital punishment. Is this claim statistically valid? Explain your reasoning.

C

7. a) Locate an article from a newspaper, periodical, or Internet site that involves a study that contains bias.

b) Briefly describe the study and its findings.

c) Describe the nature of the bias inherent in the study.

d) How has this bias affected the results of the study?

e) Suggest how the study could have eliminated the bias.

8. Inquiry/Problem Solving Do you think that the members of Parliament are a representative sample of the population? Why or why not?