2.4 – Bias and Survey Design

MDM4U
Jensen

If you conduct a survey and collect information firsthand, this is called primary data. This type of data is easy to work with because you control how it is collected.

Information obtained from similar studies conducted by OTHER researchers is called secondary data.

Part 1: Principles of Survey Design

Basic Principle #1:
A survey is not merely a collection of questions, thrown together without purpose—surveys should be designed around specific needs for information about a relevant topic.

Basic Principle #2:
Both parties to the survey have responsibilities:

- The interviewer's work must be mostly done in advance; identify relevant variables, craft questions, design the flow of the survey.
- The interviewee's task is to—having agreed to answer questions—be truthful.

Basic Principle #3:
A prime task of the interviewer at the question design stage is to help the interviewee be honest.

Part 2: Open vs. Closed Questions

1. Open Questions

- answered in respondents own words
- wide variety of possibilities
- answers sometimes difficult to interpret

Examples:

How do you feel about the salaries paid to professional athletes?

What is the most important issue for King's students?

2. Closed Questions

- respondents select from a given list of responses or the question requires an exact response
- answers are easily analyzed
- options present may bias results

Part 3: Types of Closed Questions

i) Information

Circle the appropriate response:

a) Gender: *M* *F*

b) Age: *under 14* *15 or 16*
17 or 18 *19 and over*

ii) Checklist

Which of the following sports do you enjoy watching? (check all that apply)

☐ *Basketball* ☐ *UFC*
☐ *Baseball* ☐ *Lacrosse*
☐ *Hockey* ☐ *Soccer*

iii) Rating – asks survey respondents to compare different items using a common scale. It can also be used just to rate one item using a scale.

How satisfies were you with your grade from the first unit test? (check the one that applies)

___ *Very dissatisfied*

___ *Dissatisfied*

___ *Satisfied*

___ *Very Satisfied*

Using a scale of 0 = not at all to 4 = very important, please rate the importance of each of the following aspects of service in a restuarant

	0	*1*	*2*	*3*	*4*
Speed of service	☐	☐	☐	☐	☐
Friendliness of staff	☐	☐	☐	☐	☐
Helpfulness of staff	☐	☐	☐	☐	☐
Value for money	☐	☐	☐	☐	☐
Taste of food	☐	☐	☐	☐	☐

iv) Ranking – asks survey respondents to compare a list of objects to one another by ORDERING them

When choosing a restaurant to eat at, please rank the following in order of importance from 1 to 4 where 1 is the most important to you and 4 is the least important to you

___ *Speed of Service* ___ *Ease of parking* ___ *Cleanliness* ___ *Friendliness of staff*

Part 4: Good vs. Bad Questions

Good Questions are: simple, specific, relevant, readable

Good Questions avoid: jargon, abbreviations, negatives, leading respondents, insensitivity

Example 1: What's wrong with each of the following questions?

1. Given the increasing problem of obesity amongst teenagers in North America, do you agree that King's should make physical education a mandatory class for every grade?

Leading respondents

2. Do you think the NHLPA should have agreed to the last CBA?

Abbreviations

3. Which player would you not select first in a fantasy hockey draft?

- □ Ovechkin
- □ Crosby
- □ Malkin
- □ Stamkos

Negatives, possibly jargon

Part 3: Types of Bias

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.

Sampling Bias: When the chosen sample does not accurately represent the population

Household Bias: When one type of respondent is overrepresented because groupings of different sizes are polled equally instead of proportionately

Non-response Bias: Occurs when an individual chosen for the sample can't be contacted or refuses to participate

Response/Measurement Bias: Refers to anything in the survey design that influences the responses. This includes but is not limited to:

- tendency of respondents to tailor responses to try to please the interviewer
- natural unwillingness of respondent to reveal personal facts or admit to bad behavior
- the wording of questions can influence responses

Example 2: Identifying Bias

You are the campaign manager for your best friend, Rebecca, who is running for student council Prime Minister. You have been asked to determine the overall level of support for Rebecca among the 1500 students at your school. Design a sampling method that will provide the least **sampling bias.**

Potential Solution - Plan A
To save time, you have decided that a sample of about 50 students will provide a good picture of the school's political landscape. Students have lunch periods 2, 3, or 4. By random draw from a hat, you have decided to conduct the survey in the cafeteria during period 4. The first 50 students who enter the cafeteria are given the questionnaire, and you instruct them to fill it out and return it to you before the end of lunch.

What is wrong with this scenario?

Non-response bias - some student may not complete or return the survey
Sampling bias - perhaps more seniors were let out of class early (seniors are over-represented)
- only 50 out of 1500 students were surveyed (should survey at least 10% of population)

Plan B
To fix the problems with Plan A, you have decided to provide a questionnaire to one person from each homeroom (your sample size is now 73). You can wait until the respondent finishes with the questionnaire to collect it. This will eliminate the non-response bias.

What is wrong with this scenario?

Sampling bias - still only 73 students out of 1500 (less than 10%)
Response bias - some students may just rush the survey to get through it or answer dishonestly
Household bias - some homerooms are bigger than others

Create a Plan C that is free from as much bias as possible:

Sample Answer: A stratified random sampling technique could be used to ensure a suitable sample of the student body. Students in each grade could be assigned a number. The appropriate number of students from each grade could then be selected by using a random number generator. The table below shows how a sample of 150 students could be selected to ensure that each grade is represented proportionately to its population. Interviews with each student selected would eliminate non-response bias.

10% of grade 9's – 42
10% of grade 10's – 42
10% of grade 11's – 36
10% of grade 12's – 30

Example 3: Identifying Sources of Response Bias

Consider the questionnaire below developed by Rebecca's friends. Identify examples of response bias.

Election Survey
(brought to you by the friends of Rebecca committee)

Circle the appropriate response

Gender: Male Female

Grade: 9 10 11 12

On Election Day, I intend to vote for:
Rebecca Mable Jacob

Circle what you would like:

more dances
more theme-dress days
more holidays
more fun

Brought to you by friends of Rebecca – may lead to respondents trying to please interviewer with answers

Grade 9, 10, 11, 12 – may confuse students taking classes in different levels

Rebecca – bolding the name may lead to more people choosing that name

More fun – not specific enough; won't generate any useful information