Illustrating the Effect of Leading Questions in an Introductory Statistics Course

Ulrike Genschel
Department of Statistics
Iowa State University
Introduction

- Leading questions are an example of questions with strong influence on the respondent’s choice
- Introduce easy to overlook bias into the data
- Influence of leading questions on survey responses is often discussed in the context of observational studies, survey data or the collection of data in general
Illustrating the Effect of Leading Questions in an Introductory Statistics Course

Overall Goal

• Actively engage and involve students when learning about leading questions

• Allow students to experience effect of leading questions firsthand

• Hope to increase students’ understanding and retention of the subject matter

• Use collected data on student responses later to illustrate various statistical methodology.
Idea

• Assign every student two questions to answer (yes or no)

• Each question is on a different topic:
  
  - Question 1 is on familiar topic - students likely to have already formed opinion about prior receiving question.
  
  - Question 2 is on topic that most students in the classroom are unfamiliar with.
Idea

- Question 1 is about alcohol consumption
- Question 2 is about principle of international trade which assures trade without discrimination
- Neutral and leading version of each question
- Wording (neutral versus leading) is randomized at student level for each question separately
Idea

Four Treatments:

1. Neutral Question 1 & Neutral Question 2 (NN)
2. Neutral Question 1 & Leading Question 2 (NL)
3. Leading Question 1 & Neutral Question 2 (LN)
4. Leading Question 1 & Leading Question 2 (LL)
Question 1

Neutral version:

Do you agree?

Discussing the effects of excessive alcohol consumption should be part of every high school curriculum.

• Yes
• No
Question 1

Leading version:

Do you agree?

Alcohol is linked to an estimated 5,000 deaths in people under the age of 21 each year - more than all illegal drugs combined. Discussing the dangers of excessive alcohol consumption should be part of every high school curriculum as a means to raise awareness to the dangers of alcohol consumption.

- Yes
- No
Question 2

Neutral version:

Do you agree?

China should be granted "Most Favored Nation" Status.

- Yes
- No
Question 2

Leading version:

Do you agree?

According to a recent Gallup poll, 80 percent of Americans support granting China "Most Favored Nation Status" as a means to continue important trade.

• Yes
• No
Implementation

- Students take survey “blindly”
- Part of an on-line assignment in Blackboard
- Allows for randomization of question type at the student level
- Easy retrieval of the responses to analyze data
Hypotheses

Hypothesis 1:

• Students are more likely to agree with the statement of the question when the question is phrased as a leading question.

Hypothesis 2:

• The effect of a leading question is more pronounced for the topic that students are unfamiliar with.
Survey Results and Data

Is there an association between the wording of the question (neutral vs. leading) and the proportion of respondents agreeing with the statement? Specifically, does the leading wording increase the proportion of respondents in favor of the statement?

\[ H_0: \ P_{\text{Leading}}(\text{Yes}) = P_{\text{Neutral}}(\text{Yes}) \]

versus

\[ H_a: \ P_{\text{Leading}}(\text{Yes}) > P_{\text{Neutral}}(\text{Yes}) \]
## Results Question 1

<table>
<thead>
<tr>
<th>Question 1 (Col%) (Row%)</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Neutral</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>195 (46.32%)</td>
<td>24 (63.16%)</td>
<td>219</td>
</tr>
<tr>
<td></td>
<td>(89.04%)</td>
<td>(10.96%)</td>
<td></td>
</tr>
<tr>
<td><strong>Leading</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>226 (53.68%)</td>
<td>14 (36.84%)</td>
<td>240</td>
</tr>
<tr>
<td></td>
<td>(94.17%)</td>
<td>(5.83%)</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>421</td>
<td>38</td>
<td>459</td>
</tr>
</tbody>
</table>
Results Question 1

- Pearson Chi-Square Test of Independence

\[ \chi^2_{df=1} = 3.962 \]

p-value < 0.0001
## Results Question 2

<table>
<thead>
<tr>
<th>Question 2 (Col%) (Row%)</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>43 (27.74%)</td>
<td>186 (61.18%)</td>
<td>229</td>
</tr>
<tr>
<td></td>
<td>(18.78%)</td>
<td>(81.22%)</td>
<td></td>
</tr>
<tr>
<td>Leading</td>
<td>112 (72.26%)</td>
<td>118 (38.82%)</td>
<td>230</td>
</tr>
<tr>
<td></td>
<td>(48.70%)</td>
<td>(51.53%)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>155</td>
<td>304</td>
<td>459</td>
</tr>
</tbody>
</table>
Results Question 2

- Pearson Chi-Square Test of Independence

\[ \chi^2_{df=1} = 45.925 \]

p-value = 0.0233
Utilizing Data

Mosaic Plots (descriptive)

One-sample z-test for a single proportions
CI: a single proportion
CI: single proportion – “adjusted Wald”

Two-sample z-test: difference in proportions
CI: difference in proportions

Two-way tables (Contingency Tables)
Chi-Square Test for Association
Fisher’s Exact Test (dependent on results)
Conclusions

• The questions have been tested twice and each time resulted in the same overall finding.

• Students generally seem more interested when collected data is about themselves
What is next?

Assessing effectiveness of the activity.

• Does activity improve student’s ability to identify leading questions?

• Does activity improve student’s retention of the concept of leading questions?
Thank you!