Making a Good Questionnaire

Dave Osthus  Stephanie Zimmer

Iowa State University Department of Statistics
Center of Survey Statistics and Methodology

November 1, 2010
Outline

1. Introduction
2. Nonsensitive Questions about Behavior
   - Wording of Questions
   - Diminish Impact of Forgetting on Accuracy of Survey Reports
3. Sensitive Questions about Behavior
   - Types of Questions Suitable for Sensitive Behaviors
   - Phrasing of Questions Suitable for Sensitive Behaviors
   - Reduce Sensitivity of Topic
   - Methods for Improving Reports for Sensitive Behavior
   - Validate Study
4. Attitude Questions
   - Wording of Questions
   - Structure of Questions
   - Structure of Response Options
   - Other Ways than Scales
In Statistics 521, we focus on “theory of design and estimation in sample surveys” according to the course catalog. Overall survey design looks like this:

![Survey Design from Survey Methodology (Groves et. al.)](image)

**Figure:** A survey from a process perspective.

**Figure:** Survey Design from Survey Methodology (Groves et. al.)
Design of Survey

In Statistics 521, we focus on “theory of design and estimation in sample surveys” according to the course catalog. Overall survey design looks like this:

![Diagram of survey design process](image)

**Figure 2.4** A survey from a process perspective.

**Figure:** Survey Design from Survey Methodology (Groves et. al.)
How to Ask Questions

- How do you ask questions to people?
- What considerations are there?
- How do people think when answering questions?
- Three types of questions
  - Nonsensitive Questions about Behavior
  - Sensitive Questions about Behavior
  - Attitude Questions

Definition: **Nonsensitive questions about behavior** are questions that require the respondent to remember unembarrassing, relevant information. Thus, writing good questions about nonsensitive behavior boils down to writing questions that reduce memory problems.

Examples: Shopping behavior, television viewing, number of doctor appointments, high school grades
With closed questions, include all reasonable possibilities as explicit response options

- Definition: A **closed question** requires the respondent to select an answer from among a list of options provided in the question.
- Respondents are often unwilling to offer additional information than what is specified
With closed questions, include all reasonable possibilities as explicit response options continued

**What is your race?**
- White
- Black
- Asian or Pacific Islander
- American Indian or Alaska Native
- Some Other Race

**What is your race?**
- White
- Black
- Asian Indian
- Chinese
- Japanese
- Korean
- Vietnamese
- Filipino
- Other Asian
- Native Hawaiian
- Guamanian
- Samoan
- Some Other Pacific Islander
- American Indian or Alaska Native
- Some Other Race
Make the questions as specific as possible

- Decreases the chance of multiple interpretations across respondents

- Example:
  - Not preferred: “In a typical week, how often do you usually have dessert?”
  - Preferred: “Over the last month, that is, since October 1, 2010, how many times did you have dessert in a typical week?”

- There is a balance here between question clarity and respondent demand
Use words that virtually all respondents will understand

- Much easier said than done
- Use everyday terms instead of technical ones (i.e. heart attack instead of myocardial infarction)
- Use frequency categories instead of vague quantifiers (i.e. “Once a month” instead of “Often”)
- If a technical term must be used, it should be defined in the question or somewhere in the questionnaire
Lengthen the questions by adding memory cues to improve recall

- Definition: A **memory cue** is a word, or phrase, or example that jogs a person’s memory
- Memory cues should increase the likelihood of information retrieval
- Example: In the past week, on average, how many times a day do you utilize a search engine? Examples of search engines include Google, Yahoo, and Bing.
When the events of interest are frequent but not very involving, have respondents keep a diary

- Diaries allow for frequent noting of potentially mundane events
- Breaks long reference windows up into manageable chunks of time
- Example: Food diaries, exercise diaries
To reduce telescoping errors, ask respondents to use household records

- **Definition:** **Telescoping** is the misdating of an event
- Household records help place memories of events in time
- Household record examples: Calendars, bills, financial records, emails, texts
- Example: In what month and year did you make your most recent television purchase?
To reduce telescoping errors in a longitudinal survey, use bounded recall

- **Definition:** **Bounded recall** is the act of reminding the respondent of their answers from their previous interview.

- **Example:** In our last interview on November 1, 2009, you reported having monthly doctor appointments with your cardiologist (heart doctor) due to an irregular heart beat. Since November 1, 2009, on average, how many times a month do you see a cardiologist, whether it be the same cardiologist or a different one?
If cost is a factor, consider whether proxies might be able to provide accurate information

- **Definition:** A *proxy* is a person who acts as a substitute for another person
- **Proxy Example:** Parent(s) or guardian
- **Question Example:** How many hours a day does your young child spend napping?
- **Proxies often yield more generic information than self-reports**
- **Don’t use proxies when the information is sensitive in nature**
- **Example:** Wanting to know cigarette, alcohol, or drug use of teenagers
Definition: **Sensitive questions** are questions about illegal, socially undesirable, or embarrassing behaviors that may seem intrusive to some respondents.

Examples: Drug use, alcohol consumption, tax evasion, pornography viewing, sexual promiscuity, stealing, cheating, etc.
Use open rather than closed questions for eliciting the frequency of sensitive behaviors

- Open questions have two advantages over closed questions.
  1. Closed questions inevitably lose information
  2. Closed categories provide information about the distribution of potential responses that can influence answers
Influence of closed question example

- Example: Schwarz, Hippler, Deutsh, and Strack (1985)
- Design: Two different randomized experiments embedded in a larger survey.
- Respondents asked to estimate the amount of television they watch per day.

<table>
<thead>
<tr>
<th>Low Options (n=132)</th>
<th>High Options (n=79)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responses</td>
<td>Responses</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>&lt; 0.5 hr</td>
<td>&lt; 2.5 hr</td>
</tr>
<tr>
<td>7.4%</td>
<td>62.5%</td>
</tr>
<tr>
<td>0.5 to 1 hr</td>
<td>2.5 to 3 hr</td>
</tr>
<tr>
<td>17.7%</td>
<td>23.4%</td>
</tr>
<tr>
<td>1 to 1.5 hr</td>
<td>3 to 3.5 hr</td>
</tr>
<tr>
<td>26.5%</td>
<td>7.8%</td>
</tr>
<tr>
<td>1.5 to 2 hr</td>
<td>3.5 to 4 hr</td>
</tr>
<tr>
<td>14.7%</td>
<td>4.7%</td>
</tr>
<tr>
<td>2 to 2.5 hr</td>
<td>4 to 4.5 hr</td>
</tr>
<tr>
<td>17.7%</td>
<td>1.6%</td>
</tr>
<tr>
<td>&gt; 2.5 hr</td>
<td>&gt; 4.5 hr</td>
</tr>
<tr>
<td>16.2%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
</tr>
<tr>
<td>100.0%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

- 16.2% and 37.5% of respondents reported watching more than 2.5 hours of television per day, in the low and high options groups, respectively.
Use familiar words in describing sensitive behaviors

- Example: Say “Having sex” rather than “Coitus” or “Drunk” rather than “Inebriated”
- As with nonsensitive questions, these terms are more likely to match the terms coded in the respondent’s memory
Deliberately load the question to reduce misreporting

- **Definition:** **Loading** a question means wording it in a way that invites a particular response
- Makes respondents feel alright about revealing a socially undesirable, but accurate answer
- **Four different approaches to loading questions:**
  1. **Everybody does it:** “Even the calmest parents get mad at their children sometimes. Did your children do anything in the past week that caused you to spank them?”
  2. **Assume the behavior:** “How many times during the past week did your children do something that caused you to spank them?”
  3. **Authorities recommend it:** “Many psychologists believe it is important for parents to express their pent-up frustrations. Did your children do anything in the past week that caused you to spank them?”
  4. **Reasons for doing it:** “Parents become angry because they’re tired or distracted or when their children are unusually naughty. Did your children do anything in the past week that caused you to spank them?”
Embed the sensitive question among other sensitive items to make it stand out less

- The perception of sensitivity is affected by context
- Example: Asking about shoplifting behavior after asking about armed robbery and manslaughter makes shoplifting seem less severe
Have respondents fill out a self-administered questionnaire

- Removes the interviewer from the process
- People feel diminished level of judgment
Use a randomized response technique

- Proposed by Warner in 1965
- The process via example: “Have you viewed pornography in the last seven days?”
  1. Before answering, respondents pick either a red or blue bead out of a hat, but only they know what color is it
  2. The color of the chosen bead indicates what question they answer
     - Red bead: “Have you viewed pornography in the last seven days?”
     - Blue bead: “Were you born in December?”
  3. Respondent honestly answers the question corresponding to bead color
  4. An estimate of percentage of people answering yes to the sensitive question can be calculated
Use a randomized response technique continued

Suppose these are the numbers:

- \( P(\text{Red Bead}) = P(\text{Blue Bead}) = \frac{1}{2} \)
- \( P(\text{December Birthday}) = \frac{1}{12} \)
- Proportion of “Yes” responses = \( P(\text{Yes}) = \frac{3}{10} \)
- The quantity of interest is the probability of answering yes to the sensitive questions, \( P(\text{Porn}) \)

\[
P(\text{Yes}) = P(\text{Red Bead}) \times P(\text{Porn}) + P(\text{Blue Bead}) \times P(\text{December Birthday})
\]

\[
\Rightarrow P(\text{Porn}) = \frac{P(\text{Yes}) - [P(\text{Blue Bead}) \times P(\text{December Birthday})]}{P(\text{Red Bead})}
\]

\[
= \frac{\frac{3}{10} - [(1/2) \times (1/12)]}{(1/2)}
\]

\[
= \frac{31}{60}
\]
Gather validation data

- Gain a sense of how accurate the data is
- Example: If survey questions were about recent drug use, could compare individuals responses to a urinalysis
- Done only on a subset of survey respondents
Attitude Questions

- Definition: **Attitude questions** are questions that ask about a person’s preference or opinion.
- Questions often involve response scales
- Response scales have some of their own problems
- Little problem with recall in these questions, not asking for a fact, just an opinion or feeling
Double-barreled Questions

- Definition: **Double barreled question** is a question which asks two things but only allows one answer.
- Example 1: Please agree or disagree with the following statement: cars should be faster and safer.
- Example 2: Should the government spend less money on military and more on education?
- Responses to these are difficult if not impossible to interpret.
- Be sure questions are only asking one thing.
- Then you are less worried about “what” you are measuring.
Measuring the strength of attitude

- Generally measuring direction of attitude and strength
- The use of bipolar items, forces respondent to choose one option
- Example: Should the government see to it that everyone has adequate medical care or should everyone see to his or her own medical care?
Order Matters

- Ask general questions before specific questions
- Leads to less over interpretation
- Example: Don’t ask about marital happiness then overall happiness
- When asking questions about multiple topics, start with least popular topics
Response Options

- Use closed questions - not open ended, this is just too difficult to analyze
- Bottom end of scale should be least popular option unless interview is done verbally
- Alternatives mentioned matter, if you have a scale, including “No opinion” or “Neither agree nor disagree”. These options should be included
- Exception is polling surveys or other things where you really want to know which way someone is leaning
Response Scales

- Use five to seven point scales, label every point, if using numeric scale, don’t use negative numbers.
- Too many options is overwhelming, too few lose information.
- If you really want more than 7 options, maybe use an analogue device like a thermometer, 0 to 100 (cold to very warm), respondents tend to round though.
Ranking Questions

- Sometimes a question is ranking items
- Example: Rank the following soda, pop, or carbonated beverages from 1 to 4 with 1 being favorite and 4 being least favorite. Ties are not allowed.
  __ Coke
  __ Pepsi
  __ Sprite
  __ Mountain Dew
- Only ask rank questions if respondent can see question (not phone interviews)
- Randomize order if possible (computer survey)
Check All That Apply Questions

- Avoid these types of questions
- Instead of asking, check all that apply, ask multiple questions

Wrong Way
Which of the following practices do you agree are easy to do everyday?
___ Recycle
___ Compost
___ Conserve electricity

Right Way
Do you think it is easy to the following activities everyday?
Circle yes or no.
Recycle yes no
Compost yes no
Conserve Electricity yes no
Questions?