Summary of Instrument and Question Design

Professor Ron Fricker
Naval Postgraduate School
Monterey, California
Example #1:
Poor “Instrument” Design?
Figure A.5  Votes for Buchanan in all Florida counties in 2000 presidential election, relative to the number of registered reform party voters (based on data provided by Sebago Associates, 2000).
Examples #2 & 3: Small Design Changes Can Decrease Errors

**Figure A.8** Example of two-part question with and without symbolic arrow within foveal view and resultant effects on the rate of improper navigation (Christian and Dillman, 2004).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6. After finishing school at Washington State University, where do you hope to live?</td>
<td>Eastern Washington</td>
<td>90.8%</td>
</tr>
<tr>
<td></td>
<td>Somewhere else</td>
<td>Where?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>With Symbolic Navigation Arrow</th>
<th>6. After finishing school at Washington State University, where do you hope to live?</th>
<th>93.9%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eastern Washington</td>
<td>4.2%</td>
</tr>
<tr>
<td></td>
<td>Somewhere else</td>
<td>Where?</td>
</tr>
</tbody>
</table>

Table: Examples of different branching instruction designs used in a census test and the total error rates of each during classroom experiments and a large national test in the U.S. Decennial Census (Redline and Dillman, 2005).

<table>
<thead>
<tr>
<th>Control Form: Standard Format Used in Decennial Census with Written Instructions</th>
<th>Total Classroom Error Rate</th>
<th>Total Census Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. LAST YEAR, 1999, did this person work at a job or business at any time? Yes □ No □ Go to 31</td>
<td>22.3%</td>
<td>25.8%</td>
</tr>
<tr>
<td>b. How many weeks did this person work in 1999?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prevention Method: Advanced Verbal Warning, Shift Answer Box to Right, and Larger, Bolder Font</th>
<th>Total Classroom Error Rate</th>
<th>Total Census Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. LAST YEAR, 1999, did this person work at a job or business at any time? Yes □ No □ Go to 31</td>
<td>11.3%</td>
<td>17.5%</td>
</tr>
<tr>
<td>b. How many weeks did this person work in 1999?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Detection Method: Arrows, Larger and Bolder Fonts, and Verbal Feedback</th>
<th>Total Classroom Error Rate</th>
<th>Total Census Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. LAST YEAR, 1999, did this person work at a job or business at any time? Yes □ No □ Go to 31</td>
<td>12.3%</td>
<td>21.7%</td>
</tr>
<tr>
<td>b. (If Yes) How many weeks did this person work in 1999? Count paid vacation, paid sick leave, and military</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Total error rates are shown here, including commission errors (not skipping ahead when directed) and omission errors (skipping ahead when not directed). Advanced verbal warning was used only in the classroom experiment. Statistically significant difference in total error rates between the control form and each of the others, with p ≤ .01 for both experiments.
Goals for this Lecture

• In this lecture, we will summarize some guidelines for instrument and question design
  – How to start a survey
  – General instrument guidelines
  – General question guidelines
  – General response scale guidelines

• As with any list of general rules, they’re sometimes wrong
  – Use these guidelines as a starting point, not as dogma that must be rigidly followed
Good Design Helps Minimize Nonresponse and Measurement Error

• Instrument should be easy to understand, to navigate, and to complete
  – Encourage response

• Questions must be:
  – Clear
  – Unambiguous
  – Minimize possible bias
    • E.g., questions with socially desirable answers

• The harder it is for the respondents, the higher the survey nonresponse rate
Make the Survey User-friendly!

- Make the instrument user-friendly and allow for flexibility
  - Make it simple to read, follow, and complete
  - Let respondents skip around, go back and change responses
  - Good electronic survey design:
    - Don’t force answers unless absolutely necessary
      - But do remind respondents if they fail to answer
    - Include response checks as much as possible
    - Automate skip patterns
- Invest time in instrument design to minimize respondent burden
  - Payoff will be in improved response rates
How to Start the Survey

• Open with easy questions all respondents can answer with little effort
  – Initial questions should also be non-threatening
• Don’t start with knowledge or awareness questions
• First questions should be directly related to the topic as described in the introduction or advance/cover letter
  – Make it immediately relevant to the respondent
Keep the Survey as Short as Possible

• Minimize total number of questions
  – Remember the Mark Twain quote: “I'm sorry this letter is so long, but I did not have time to make it shorter.”
  – Bottom line: *If you aren’t going to analyze it, don’t ask it!*
    • Corollary: If it doesn’t help answer the research objective, don’t ask it

• Ensure respondents have ability to respond via chosen survey modes
  – Web-based survey: do all potential respondents have Internet access?
    • Else, could result in *selection bias*
  – Allow for alternate response modes as necessary
Physical Format Checklist

• Number all questions sequentially
• Use large, clear type
  – Make it easy to read
• Use ‘white space’
  – Don’t crowd questions and text
  – Place more blank space between questions than between subcomponents of questions
• Put special instructions on questionnaire as needed, next to (or within) question
• Clearly distinguish directions from questions
  – Use different fonts, white space, boxes, etc.
Physical Format, Continued

- Don’t split questions across pages/screens
  - If necessary (e.g., question requires 1.5 pages), restate question and response categories on next page
- If possible, list response categories vertically instead of horizontally
- Be consistent with
  - Direction of response categories
  - Placement of response categories
Organizing the Questions

- Segment questions by topic
- Ask about related topics together
- Salient questions take precedence over less salient ones
- Ask recall backwards in time
- Use transitions when changing topics – give a sense of progress through the questionnaire
- Leave objectionable questions (e.g., income) for the end
- Put demographic questions at the end
• Questions must be:
  – Easily and consistently understood
  – Easy to answer:
    • Only ask one question at a time
    • Break complicated questions/concepts up into a series of
      simpler questions
• Respondents have a common understanding of
  – The meaning of each question
  – What constitutes an adequate answer
• Respondents also must:
  – Have access to information needed to answer
  – Be willing to provide answers called for by questions
Eliciting Correct Information Not As Simple As One Might Expect

A simple model of the cognitive process in answering questions:

Figure 7.1 A simple model of the survey response process.
• **Comprehension** includes such processes as
  – Attending to the question, and instructions;
  – Assigning a meaning to the surface form of the question; and,
  – Inferring the question’s point

• **Interpreting the question entails**
  – Parsing the question
  – Assigning meanings to the key elements
  – Inferring the purpose behind the question
  – Determining the boundaries and potential overlap between permissible answers
Retrieval

- **Retrieval** is the process of recalling information relevant to answering the question from long-term memory (or from wherever the information is physically stored).

- Successful retrieval can depend on:
  - Nature of the events in question
  - How long ago the events took place
  - The number and richness of the “cues” that initiate the retrieval process
Estimation and Judgment

• **Estimation and judgment** are the processes of combining or supplementing what the respondent has retrieved.

• Essentially, the respondent must figure out how to quantify or rate the information being retrieved to appropriately answer the question.

• May also have to literally estimate their answer from imperfect memory /records of events.
• **Reporting** is the process of selecting and communicating an answer, which includes
  – Mapping the answer onto the question’s response options
  – Altering the answer for consistency with prior answers, perceived acceptability, etc.

• How respondents report their answer will depend in part on the fit between the information they retrieve and constraints imposed by the question
Problems in Answering Survey Questions

1. Failure to encode information sought
2. Misinterpretation of the question
3. Forgetting and other memory problems
4. Flawed judgment or estimation strategies
5. Problems in formatting answer
6. More or less deliberate misreporting
7. Failure to follow instructions
• People are notoriously poor eyewitnesses
  – Similarly, they often miss key details of many things observed / experienced
• A.F. Smith study compared what people said they ate in a survey to what was recorded in detailed food diaries the respondents kept
  – He concluded, “dietary [survey] reports…consist in large part of individuals’ guesses about what they probably ate.”
• Conclusion: People cannot report what they have not mentally “encoded”
• Respondents are unlikely to provide accurate answers to misinterpreted questions
• Seven types of comprehension problems:
  – Grammatical ambiguity
  – Excessive complexity
  – Faulty presupposition
  – Vague concepts
  – Vague quantifiers
  – Unfamiliar terms
  – False inferences
Figure 7.2 Recall accuracy for types of personal information. (Source: Tourangeau, Rips, and Rasinski, 2000.)
### Factors Affecting Recall

#### Table 7.1. Summary of Factors Affecting Recall

<table>
<thead>
<tr>
<th>Variable</th>
<th>Finding</th>
<th>Implication for Survey Design</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Characteristics of Event</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time of occurrence</td>
<td>Events that happened long ago are harder to recall</td>
<td>Shorten the reference period</td>
</tr>
<tr>
<td>Proximity to temporal boundaries</td>
<td>Events near significant temporal boundaries are easier to recall</td>
<td>Use personal landmarks, life events calendars to promote recall</td>
</tr>
<tr>
<td>Distinctiveness</td>
<td>Distinctive events are easier to recall</td>
<td>Tailor the length of the reference period to the target events; use multiple cues to single out individual events</td>
</tr>
<tr>
<td><strong>Importance, emotional impact</strong></td>
<td>Important, emotionally involving events are easier to recall</td>
<td>Tailor the length of reference period to the properties of the target events</td>
</tr>
<tr>
<td><strong>Question Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recall order</td>
<td>Backward search may promote fuller recall</td>
<td>Not clear whether backward recall better in surveys</td>
</tr>
<tr>
<td>Number and type of cues</td>
<td>Multiple cues typically better than single cue;</td>
<td>Provide multiple cues; use decomposition</td>
</tr>
<tr>
<td></td>
<td>cues about the type of event (what) better than cues about participants or location (who or where), which are better than cues about (when)</td>
<td></td>
</tr>
<tr>
<td>Time on task</td>
<td>Taking more time improves recall</td>
<td>Use longer introductions to questions; slow the pace of the interview</td>
</tr>
</tbody>
</table>

24
• Once respondents have estimated an answer, they must translate it into an acceptable answer
• More or less difficult depending on the response format:
  – open-ended
  – closed-ended with ordered response scale
  – closed-ended with categorical response options
• Good design can mitigate some problems, but some will likely remain
Motivated Misreporting

• Potential issue with sensitive questions
  – It may be easier for a respondent to under-report something than to refuse to answer the question
  – See techniques in lecture on question design, such as
    • “Forgiving” wording
    • Randomized response technique
Question Design Guidelines

• *Most* survey question design isn’t rocket science
  – It’s mainly common sense combined with careful pre-testing of questions and responses

• Keep questions as simple and direct as possible
  – Break complex queries into a series of questions

• Avoid ambiguous terms and questions like the plague
  – Corollary: Carefully and clearly define complex and technical terms

• Only ask open-ended questions if necessary
  – Don’t ask respondents to do your work for you
  – Exception: As appropriate, give respondents a chance to provide input not otherwise asked for
Disadvantages of Closed Questions

- Categories may be leading to respondents
- May make it too easy to answer without thinking
- May limit spontaneity
- Not best when
  - Asking for frequency of sensitive behaviors
  - There are numerous possible responses
Open-ended Questions

• Useful for allowing respondents to provide additional data not asked for in the survey
  – Can mitigate possible frustration w/ survey

• Also good for
  – Understanding reasoning behind other answers
  – When possible responses exceeds reasonable list
  – May be easiest way to collect information about complicated phenomenon/issue/problem

• But, use sparingly:
  – Generally hard to analyze
    • New software making this less of a problem
  – If overused, puts large burden on respondents
Don’t Forget “Don’t Knows”

• Lack can frustrate some respondents
  – If respondents forced to just skip answering a question (item nonresponse) can make analysis more difficult

• Solutions:
  – Ask a screening question first
    • Appropriately allows instrument to build in “skip patterns” for respondents to skip over questions that do not apply
  – Include a “no opinion” or “don’t know” response
  – Force respondent to answer the question (tricky to do well)
A Bit About Response Scales

• There are many possible ordinal scales:

![Three Scales Diagram]

- Good
- Not Good
- Good
- Fair
- Poor
- Excellent
- Very Good
- Good
- Fair
- Poor

**Figure 3.3. Three Scales**

• In structuring response scales, must balance:
  - Analytical needs (precision of information required)
  - Ease of respondent recall and answer

The “Likert” Scale is Very Common

• Wikipedia: “A Likert scale (pronounced 'lick-ert') is a type of psychometric response scale often used in questionnaires, and is the most widely used scale in survey research. When responding to a Likert questionnaire item, respondents specify their level of agreement to a statement.”

• A five-point scale is frequently used, but it can vary

• Example:

  Ice cream is good for breakfast
  (1) Strongly disagree
  (2) Disagree
  (3) Neither agree nor disagree
  (4) Agree
  (5) Strongly agree

http://en.wikipedia.org/wiki/Likert_scale
Constructing Response Scales

• Number of points in scale should be determined by how you plan to use the data
• Respondents can generally only remember a maximum of 5 responses unless visual cues are used
• With scales with few points, every scale can be labeled
  – In longer scales, only label midpoint and endpoints
• Graphic images such as thermometers and ladders can be effective
Use of Graphical Scales

Which face comes closest to expressing how you feel about _______?

**Figure A.1. Using Faces to Measure Feelings**

<table>
<thead>
<tr>
<th>WARM</th>
<th>100</th>
<th>Very warm or favorable feeling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>85</td>
<td>Good warm or favorable feeling</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>Fairly warm or favorable feeling</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>A bit more warm or favorable than cold feeling</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>No feeling at all</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>A bit more cold or unfavorable feeling</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>Fairly cold or unfavorable feeling</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>Quite cold or unfavorable feeling</td>
</tr>
<tr>
<td>COLD</td>
<td>0</td>
<td>Very cold or unfavorable feeling</td>
</tr>
</tbody>
</table>

Where would you put _______ on the feeling thermometer?

Specific Response Scale Guidelines (1)

• Clearly label the response scale(s)
• Be consistent with the scale labels and ordering throughout the instrument
• Keep the number of response categories to the necessary minimum:
  – At what level of detail can respondents easily recall/respond?
  – How finely can respondents meaningfully answer?
  – What is necessary for the analysis?
• If “n/a” response needed/appropriate, visually distance it from the rest of the scale
Specific Response Scale Guidelines (2)

• Response categories should be consistent with the question
  – In the absence of another need/requirement, default to five-point Likert scale
• Categories must be exhaustive, including every possible answer
• Categories must be mutually exclusive (no overlap)
• If appropriate, include a “don’t know” category
Don’t Forget to Collect Relevant Covariate Data

• Usually it is important to collect covariate data for the analysis
  – Depends on the research objectives, of course
• For example, demographic data
  – E.g., for individuals: age, gender, race, marital status, military service, rank, etc.
  – E.g., for organizations: size, type, mission, etc.
• Usually these are important categories that you want to conduct the analysis by
Take Away: Good Survey Design is a Science and an Art

For a good survey:
✓ You must ask the right questions
✓ Respondents must understand your questions
✓ Respondents must know the answers
✓ Respondents must be willing and able to tell you the answers

And remember, always pretest!
What We Have Covered

• In this lecture, we have summarized some guidelines for instrument and question design
  – How to start a survey
  – General instrument guidelines
  – General question guidelines
  – General response scale guidelines

• Remember, as with any list of general rules, they’re sometimes wrong
  – Use these guidelines as a starting point, not as dogma that must be rigidly followed.