Always remember that questionnaire data are limited by error due to *self-reporting*.

**Prevalence of Current Smoking**

<table>
<thead>
<tr>
<th>Region</th>
<th>18-24</th>
<th>25-44</th>
<th>45-64</th>
<th>65+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northeast</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education</th>
<th>&lt;12 years</th>
<th>12 years (H.S.)</th>
<th>13-15 years</th>
<th>16+ years</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Estimated Prevalence of Current Smoking, Determined through Self-Report**

<table>
<thead>
<tr>
<th>Region</th>
<th>18-24</th>
<th>25-44</th>
<th>45-64</th>
<th>65+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northeast</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education</th>
<th>&lt;12 years</th>
<th>12 years (H.S.)</th>
<th>13-15 years</th>
<th>16+ years</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Questionnaire development approach
(See Aday, L., & Cornelius, L. (2006). Designing and Conducting Health Surveys, Wiley)

I. Determine Analytic Objectives
   • What types of data will answer the research question?

II. Develop general concepts to be covered
   • List areas to be covered by questions

III. Translate concepts into questions

IV. “Appraise” questions for common pitfalls

V. Evaluate questions empirically
Questionnaire development approach
(See Aday, L., & Cornelius, L. (2006). Designing and Conducting Health Surveys, Wiley)

I. Determine Analytic Objectives
   • What types of data will answer the research question?
II. Develop general concepts to be covered
   • List areas to be covered by questions
III. Translate concepts into questions
IV. “Appraise” questions for common pitfalls
V. Evaluate questions empirically
Appraise questions for common pitfalls

- First, consider Administration Mode:
  - Problems tend to be specific to choice of:
    - Interviewer administration
      - Telephone
      - In-person
    - Self-Administration
      - Mailed paper
      - Internet
      - Personal Digital Assistant (PDA)/Smartphone

Sources of Response Error: Tourangeau (1984) cognitive model

- Encoding of question (understanding it)
  - Have you ever received care from an oral surgeon?
- Retrieval of information (knowing/remembering)
  - How many times have you ridden in a passenger airplane?
- Decision and judgment processes (truth, adequacy)
  - How many sex partners have you had in the past 12 months?
- Response (matching internal representation to given categories)
  - Would you say your health is excellent, very good, good, fair, or poor?

Evaluation: How do we find questionnaire problems?

  - Question Appraisal System (Willis & Lessler, 1999):
Lack of Clarity: Difficult-to-Understand Questions

Long/Convoluted Phrasing:

- The last time that you were seen by a doctor, nurse, or other health professional, as part of a regular medical check-up, did you receive any tests specifically designed to diagnose the presence of certain types of cancer?

- Typical response = “What?”

⇒ Especially for interviewer-administration: DECOMPOSE question into concepts -- ask more, but simpler questions, with use of skips

“Decomposition” into simpler phrasing

- When did you last see a doctor, nurse, or other health professional, to get a regular medical check-up?

- During that visit, did you receive any tests that check for cancer?

- What types of cancer were you checked for?

⇒ Doesn’t solve problem of respondents not knowing the answer, but makes the question more understandable.

⇒ A VERY common problem -- and a very common “fix”
Lack of clarity: Terms/phrases are difficult to understand

Complex/Unfamiliar Terminology:

- Were you seen on an inpatient or outpatient basis?
- Have you ever had a colonoscopy or sigmoidoscopy?

Better to use “simple language”:

- Did you stay overnight at the hospital?
- (Use explanation of what the medical test entails)

Question Clarity/Vagueness

- Many questions that use “simple language” are variably interpreted:
  - Have you ever been a regular smoker?
  - Does anyone in your family now have a car?
  - Do you think that headaches can be effectively treated?

Lack of Clarity/Vagueness

Have you had your blood tested for the AIDS virus?

- Unclear: a) “Did I take the initiative in deciding to have my blood tested?” versus b) “Was it tested as part of any type of blood test?”
- If the issue of interest is the act of testing, simply ask “As far as you know, has your blood ever been tested for the AIDS virus?”
Retrieval problem:
Respondent doesn’t know the answer

Estimate the number of your women patients with whom you discussed enrollment in a cancer TREATMENT trial in the LAST 12 MONTHS:

<table>
<thead>
<tr>
<th></th>
<th>ALL WOMEN</th>
<th>ASIAN AMERICAN WOMEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Cancer treatment trials</td>
<td>____</td>
<td>____</td>
</tr>
<tr>
<td>Breast Cancer Treatment Trials</td>
<td>____</td>
<td>____</td>
</tr>
</tbody>
</table>

Logical/Cultural Problems

- Questions that simply don’t make sense in case of at least some respondents
  - Sometimes due to cultural issues—“Have you ever switched from a stronger to a lighter cigarette”
- Problems exist separately from respondent’s interpretation, recall, decision processes
- Avoidance, resolution requires clear understanding of environment, culture, ‘respondent reality’

Formatting Problems

- Self Administered:
  - Especially for mail survey, DON’T CROWD QUESTIONS ONTO PAGES (better to have more pages, “open” layout)
  - If paper (not computer): MINIMIZE use of skip patterns - respondents get confused very easily, make errors
Formatting Problems (Cont’d)

• Interviewer Administered:
  – Remember that the respondent doesn’t read along—CANNOT be designed like self-administered form
  How many glasses (8 oz) of milk (whole, 2%, or skim milk) did you drink yesterday?
  – Need to use HAND/SHOW CARDS for long lists
  – Response categories implied should match those on form:

Format problem: Mis-match of question, answer categories

*How do you feel about your present weight? (Do not read)*:
  ___ Overweight
  ___ Underweight
  ___ About right

*In the past ten years, how many times have you had a headache severe enough to cause you to stay in bed for more than half a day? (Do not read)*:
  ___ Never
  ___ 1-5 times
  ___ 6-10 times
  ___ more than 10 times
Excessive Length

- PROBABLY THE SINGLE MOST SERIOUS PROBLEM IN FEDERAL SURVEYS
- Length will increase survey costs, decrease response rate
- Length/burden will lead to response error in unmeasurable ways
  - Interviewers read faster, less carefully (get it done!)
  - Respondents answer less carefully as time goes on
- If possible, don’t go over (average) 30 minutes for face-face, 15 minutes for phone survey, 15 minutes for Web survey

Evaluation: How do we find questionnaire problems?

- Cognitive interviewing: Search for underlying problems

Manual available at:

Book:

The cognitive testing process in a nutshell

- Develop a questionnaire or material (advance letter, etc.) to be evaluated
- Recruit (paid) members of the targeted population (e.g., recipients of home loans, people without employment, cancer survivors)
- Conduct one-on-one interviews, in “laboratory” or other location:
  - Home
  - Homeless shelter
  - Health clinic
  - Elderly center
- Use both Think-Aloud and Verbal Probing techniques
“Classic” verbal probes

**Comprehension probe:** What does the term “dental sealant” mean to you?

**Paraphrase:** Can you repeat the question in your own words?

**Confidence judgment:** How sure are you that your health insurance covers...

**Recall probe:** How do you know that you went to the dentist 3 times...?

**‘General’ probe:** How did you arrive at that answer?

Tested (“classic”) question: Pain in the abdomen

“In the last year have you been bothered by pain in the abdomen?”

What probes make sense here?

- What time period are you thinking about, exactly?
- What does “bothered by pain” mean to you?
- Where is your “abdomen?” ➔
Using cognitive interviews to detect question wording problems

VERSION 1 (No filter)
On a typical day, how much time do you spend doing strenuous physical activities such as lifting, pushing, or pulling?

___ None ___ Less than 1 hour ___ 1-4 hours ___ 5 + hours

VERSION 2 (Filtered)
On a typical day, do you spend any time doing strenuous physical activities such as lifting, pushing, or pulling?
IF YES: Read Version 1

Survey experiment results: Reporting of strenuous physical activity

On a typical day, how much time do you spend doing strenuous physical activities such as lifting, pushing, or pulling?

<table>
<thead>
<tr>
<th></th>
<th>FIELD PRETEST (n=78)</th>
<th>WOMEN’S HEALTH (n=191)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>32%</td>
<td>4%</td>
</tr>
<tr>
<td>&lt; 1</td>
<td>32%</td>
<td>42%</td>
</tr>
<tr>
<td>1-4</td>
<td>35%</td>
<td>50%</td>
</tr>
<tr>
<td>5+</td>
<td>0%</td>
<td>4%</td>
</tr>
</tbody>
</table>

fil  i  filt er

version

version

version

version
Psychometric Approaches

Classical Psychometrics

• Test-retest reliability
• Internal Consistency Reliability (Cronbach alpha)
• Factor Analysis

Modern Measurement Approaches

• Item Response Theory (IRT):
  - To what extent does each item measure the level of the underlying construct (concept)

• Differential Item Functioning (DIF)
  - Does an item reflect the level of the construct variably, for different subgroups (gender, race...)?

Psychometric Approaches: Limitations

• Psychometric approaches generally assume:
  - Use of latent construct (unobservable, underlying)
  - Use of multi-item scale (more than one item to measure construct)
• NOT applicable for single-item, behavioral / observed variables that we often measure!
  - "In the past month, have you smoked a cigarette, even one puff?"

(IRT) Item Information Curve: Multi-Item Scale asking about racial/ethnic discrimination
Useful Internet Resources

- 1) US Census Bureau guide to self-administration:

- 2) NCI resources:
  - a) The Question Appraisal System (Full version):
  - b) Cognitive Interviewing: A How-To Guide:
  - c) An Introduction to Modern Measurement Theory:

In closing…

“The uncreative mind can spot wrong answers, but it takes a very creative mind to spot wrong questions.”

- Anthony Jay