Methods for Collecting Survey Data

(Click icon for audio)

Dr. Michael R. Hyman, NMSU
There is no best form of survey; each has advantages and disadvantages
Communicating with Respondents

• Personal interviews
  – Door-to-door
  – Shopping mall intercepts
• Telephone interviews
• Self-administered questionnaires
  – Mail
  – Internet (Web and e-mail)
Factors to Consider When Choosing a Collection Method

- Available budget
- How quickly data needed
- Complexity/structure/length of questionnaire
- Need for stimuli exposure
- Sampling precision
- Quality of data
- Respondent interest in topic
- Ease of obtaining respondent cooperation
- Incidence rate
Personal Interviews
Door-to-Door Interviews

• Speed of data collection
  – Moderate to fast
• Geographical flexibility
  – Limited to moderate
• Respondent cooperation
  – Excellent
• Versatility of questioning
  – Quite versatile
Door-to-Door Interviews

- Questionnaire length
  - Long
- Item non-response
  - Low
- Possibility of respondent misunderstanding
  - Lowest
Door-to-Door Interviews

- Degree of interviewer influence of answer
  - High
- Supervision of interviewers
  - Moderate
- Anonymity of respondent
  - Low
Door-to-Door Interviews

• Ease of call back or follow-up
  – Difficult
• Cost
  – Highest
• Special features
  – Visual materials may be shown or demonstrated; extended probing possible
Mall Intercept Interviews

- Speed of data collection
  - Fast
- Geographical flexibility
  - Confined, urban bias
- Respondent cooperation
  - Moderate to low
  - Versatility of questioning
  - Extremely versatile
Mall Intercept Interviews

• Speed of data collection
  – Fast
• Geographical flexibility
  – Confined, urban bias
• Respondent cooperation
  – Moderate to low
  – Versatility of questioning
  – Extremely versatile
Mall Intercept Interviews

- Questionnaire length
  - Moderate to long
- Item non-response
  - Medium
- Possibility of respondent misunderstanding
  - Lowest
Mall Intercept Interviews

- Degree of interviewer influence of answers
  - Highest
- Supervision of interviewers
  - Moderate to high
- Anonymity of respondent
  - Low
Mall Intercept Interviews

• Ease of call back or follow-up
  – Difficult
• Cost
  – Moderate to high
• Special features
  – Taste test, viewing of TV commercials possible
Telephone Surveys

- Central location interviewing
- Computer-assisted telephone interviewing
- Computerized voice-activated interviews
Telephone Surveys

- Speed of data collection
  - Very fast
- Geographical flexibility
  - High
- Respondent cooperation
  - Good
- Versatility of questioning
  - Moderate
Telephone Surveys

- Questionnaire length
  - Moderate
- Item non-response
  - Medium
- Possibility of misunderstanding
  - Average
- Degree of interviewer influence of answer
  - Moderate
Telephone Surveys

• Supervision of interviewers
  – High, especially with central location WATS interviewing

• Anonymity of respondent
  – Moderate

• Ease of call back or follow-up
  – Easy
Telephone Surveys

- Cost
  - Low to moderate
- Special features
  - Fieldwork and supervision of data collection are simplified; quite adaptable to computer technology
Self-Administered Questionnaires Can Be Either Printed or Electronic

- Paper Questionnaires
  - Mail
  - In-person drop-off
  - Inserts
  - Fax

- Electronic Questionnaires
  - E-mail
  - Internet Web site
  - Interactive kiosk
Mail Surveys

- One Shot
- Mail Panel
Mail Surveys

• Speed of data collection
  – Researcher has no control over return of questionnaire; slow

• Geographical flexibility
  – High

• Respondent cooperation
  – Moderate--poorly designed questionnaire will have low response rate
Mail Surveys

• Versatility of questioning
  – Highly standardized format
• Questionnaire length
  – Varies depending on incentive
• Item non-response
  – High
Mail Surveys

• Possibility of respondent misunderstanding
  – Highest--no interviewer present for clarification

• Degree of interviewer influence of answer
  – None--interviewer absent

• Supervision of interviewers
  – Not applicable
Mail Surveys

• Anonymity of respondent
  – High
• Ease of call back or follow-up
  – Easy, but takes time
• Cost
  – Lowest
Internet Surveys

• Self-administered questionnaire posted on a Web site
• Respondents provide answers to questions displayed online by highlighting a phrase, clicking an icon, or keying in an answer
U.S. Mint Customer Survey

We want to know how to improve the site to serve you better. So, from time to time, we will be asking questions that will help us determine how to make the site more useful.

1. How many times have you visited the U.S. Mint Web site?
   -choose-

2. How did you hear about our Web site?
   -choose-

   2a. If other, how?

3. Which response best describes you?
   -choose-

   3a. If other, who?

4. Overall, how are we doing?
   (Excellent, Very Good, Average, Fair, Poor)

   4a. Please tell us why:

5. Have you purchased anything on the Mint Web site?
   Yes, No

6. How would you describe your online shopping experience?
Internet Surveys

• Speed of data collection
  – Instantaneous
• Cost effective
• Geographic flexibility
  – Worldwide
• Visual and interactive
Internet Surveys

- Versatility of questioning
  - Extremely versatile
- Questionnaire length
  - Individualized base on respondent answers
  - Longer questionnaires with panel samples
- Item non-response
  - Software can assure none
Internet Surveys

• Possibility for respondent misunderstanding
  – High
• Interviewer influence of answers
  – None
• Supervision of interviewers not required
Internet Surveys

• Anonymity of Respondent
  – Respondent can be anonymous or known

• Ease of Callback or Follow-up
  – Difficult unless e-mail address is known

• Special Features
  – Allows graphics and streaming media
Internet Surveys: Sample Representativeness

- Subject to self-selection bias
- Some people without Internet access
- Some people lack powerful PCs with high-speed Internet connections
- Some respondents relatively unsophisticated computer users
- When opt-in not used, cooperation is low
E-Mail Surveys

• Speed of data collection
  – Instantaneous
• Geographic flexibility
  – Worldwide
• Cheaper distribution and processing costs
E-Mail Surveys

• Flexible, but
  – Extensive differences in capabilities of respondents’ PCs and e-mail software limit types of questions and layouts

• E-mails are not secure and “eavesdropping” can possibly occur

• Respondent cooperation
  – Varies if e-mail is seen as “spam”
Summaries of Relative Advantages of Disadvantages of Different Data Collection Methods
## Comparative Indices of Direct Costs per Completed Interview (Including Travel and Telephone Charges, Interviewer Compensation, Training, and Direct Supervision Expenses)

<table>
<thead>
<tr>
<th>Data Collection Method</th>
<th>Index of Cost&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mail survey (costs depend on return rate, incentives, and follow-up procedure)</td>
<td>0.3–0.8</td>
</tr>
<tr>
<td>2. Telephone interviews</td>
<td>0.5–0.8</td>
</tr>
<tr>
<td>a. Seven-minute interview with head of household in metropolitan area</td>
<td>0.5–0.8</td>
</tr>
<tr>
<td>b. Fifteen-minute interview with small segment of national population from a central station</td>
<td>1.3–1.7</td>
</tr>
<tr>
<td>3. Personal interviews</td>
<td>1.5–1.8</td>
</tr>
<tr>
<td>a. Ten-minute personal interview in middle-class suburban area (1 call-back and 10 percent validation)</td>
<td>1.5–1.8</td>
</tr>
<tr>
<td>b. Forty- to 60-minute interview of national probability sample (1 call-back and 10 percent validation)</td>
<td>2.5–3.5</td>
</tr>
<tr>
<td>c. Executive (VIP) interviews</td>
<td>4.0–15.0+</td>
</tr>
</tbody>
</table>

<sup>a</sup> In 1997, an index value of 1.0 corresponded to a cost of $20.00.
# Advantages and Disadvantages of Typical Survey Methods

<table>
<thead>
<tr>
<th></th>
<th>Door-to-Do</th>
<th>Mail Intercept</th>
<th>Telephone</th>
<th>Mail</th>
<th>Internet</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Personal Interview</td>
<td>Personal Interview</td>
<td>Interview</td>
<td>Survey</td>
<td>Survey</td>
</tr>
<tr>
<td>Speed of data collection</td>
<td>Moderate to fast</td>
<td>Fast</td>
<td>Very fast</td>
<td>Slow; researcher has no control over return of questionnaire</td>
<td>Instantaneous; 24/7</td>
</tr>
<tr>
<td>Geographic flexibility</td>
<td>Limited to moderate</td>
<td>Confined, possible urban bias</td>
<td>High</td>
<td>High</td>
<td>High (worldwide)</td>
</tr>
<tr>
<td>Respondent cooperation</td>
<td>Excellent</td>
<td>Moderate to low</td>
<td>Good</td>
<td>Moderate; poorly designed questionnaire will have low response rate</td>
<td>Varies depending on Web site; high from consumer panels</td>
</tr>
<tr>
<td>Versatility of questioning</td>
<td>Quite versatile</td>
<td>Extremely versatile</td>
<td>Moderate</td>
<td>Not versatile; requires highly standardized format</td>
<td>Extremely versatile</td>
</tr>
<tr>
<td>Questionnaire length</td>
<td>Long</td>
<td>Moderate to long</td>
<td>Moderate</td>
<td>Varies depending on incentive</td>
<td>Moderate; length customized based on answers</td>
</tr>
<tr>
<td>Item non-response rate</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>Software can assure none</td>
</tr>
<tr>
<td>Possibility for respondent misunderstanding</td>
<td>Low</td>
<td>Low</td>
<td>Average</td>
<td>High; no interviewer present for clarification</td>
<td>High</td>
</tr>
<tr>
<td>Degree of interviewer influence on answers</td>
<td>High</td>
<td>High</td>
<td>Moderate</td>
<td>None; interviewer absent</td>
<td>None</td>
</tr>
<tr>
<td>Supervision of interviewers</td>
<td>Moderate</td>
<td>Moderate to high</td>
<td>High, especially with central location WATS interviewing</td>
<td>Not applicable</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Anonymity of respondent</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>High</td>
<td>Respondent can be either anonymous or known</td>
</tr>
<tr>
<td>Ease of callback or follow-up</td>
<td>Difficult</td>
<td>Difficult</td>
<td>Easy</td>
<td>Easy, but takes time</td>
<td>Difficult, unless e-mail address is known</td>
</tr>
<tr>
<td>Cost</td>
<td>Lowest</td>
<td>Moderate to high</td>
<td>Low to moderate</td>
<td>Respondent may answer questions at own convenience; has time to reflect on answers</td>
<td></td>
</tr>
<tr>
<td>Special features</td>
<td>Visual materials may be shown or demonstrated; extended probing possible</td>
<td>Taste tests, viewing of TV commercials possible</td>
<td>Fieldwork and supervision of data collection are simplified; quite adaptable to computer technology</td>
<td>Streaming media software allows use of graphics and animation</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The emphasis is on typical surveys. For example, an elaborate mail survey may be far more expensive than a short personal interview, but this generally is not the case.
<table>
<thead>
<tr>
<th>Criteria</th>
<th>Telephone Interviews</th>
<th>CATI</th>
<th>In-Home</th>
<th>Mall Intercept</th>
<th>CAPI</th>
<th>Mail Surveys</th>
<th>Mail Panels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flexibility of Data Collection</td>
<td>Moderate to High</td>
<td>Moderate to High</td>
<td>High</td>
<td>High</td>
<td>Moderate to High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Diversity of Questions</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Use of Physical Stimuli</td>
<td>Low</td>
<td>Low</td>
<td>Moderate to High</td>
<td>High</td>
<td>High</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Sample Control</td>
<td>Moderate to High</td>
<td>Moderate to High</td>
<td>Potentially High</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Low</td>
<td>Moderate to High</td>
</tr>
<tr>
<td>Control of Data Collection Environment</td>
<td>Moderate to High</td>
<td>Moderate to High</td>
<td>Moderate to High</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Control of Field Force</td>
<td>Moderate to High</td>
<td>Moderate to High</td>
<td>Low</td>
<td>Moderate</td>
<td>Moderate</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Quantity of Data</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Moderate</td>
<td>High</td>
</tr>
<tr>
<td>Response Rate</td>
<td>Moderate to High</td>
<td>Moderate</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Perceived Anonymity of the Respondent</td>
<td>Moderate to High</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Social Desirability</td>
<td>Moderate to High</td>
<td>Moderate</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Obtaining Sensitive Information</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low to Moderate</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Potential for Interviewer Bias</td>
<td>Moderate to High</td>
<td>Moderate</td>
<td>High</td>
<td>Low</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Speed</td>
<td>High</td>
<td>High</td>
<td>Moderate to High</td>
<td>Moderate to High</td>
<td>Low</td>
<td>Low to Moderate</td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td>Moderate to High</td>
<td>Moderate</td>
<td>High</td>
<td>Moderate to High</td>
<td>Moderate to High</td>
<td>Low</td>
<td>Low to Moderate</td>
</tr>
</tbody>
</table>
# Comparison of mail, telephone, and personal interview surveys

<table>
<thead>
<tr>
<th></th>
<th>Mail Survey</th>
<th>Telephone Survey</th>
<th>Personal Interview Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost (assuming a good response rate)</td>
<td>often lowest</td>
<td>usually in-between</td>
<td>usually highest</td>
</tr>
<tr>
<td>Ability to probe</td>
<td>no personal contact or observation</td>
<td>some chance for gathering additional data through elaboration on questions, but no personal observation</td>
<td>greatest opportunity for observation, building rapport, and additional probing</td>
</tr>
<tr>
<td>Respondent ability to complete at own convenience</td>
<td>yes</td>
<td>no</td>
<td>perhaps, if interview time is prearranged with respondent</td>
</tr>
<tr>
<td>Interviewer bias</td>
<td>no chance</td>
<td>some, perhaps due to voice inflection, etc.</td>
<td>greatest chance</td>
</tr>
<tr>
<td>Ability to decide who actually responds to the questionnaire within a household</td>
<td>least</td>
<td>some</td>
<td>greatest</td>
</tr>
<tr>
<td>Sampling problems</td>
<td>up-to-date, accurate mailing list and low response rates</td>
<td>up-to-date, accurate phone subscriber list, unlisted numbers, no phones, refusals</td>
<td>not-at-homes and refusals</td>
</tr>
<tr>
<td></td>
<td>Mail Survey</td>
<td>Telephone Survey</td>
<td>Personal Interview Survey</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Impersonality</td>
<td>greatest</td>
<td>some due to lack of face-to-face contact</td>
<td>least</td>
</tr>
<tr>
<td>Complex questions</td>
<td>least suitable</td>
<td>somewhat suitable</td>
<td>most suitable</td>
</tr>
<tr>
<td>Visual aids in survey</td>
<td>little opportunity</td>
<td>no opportunity</td>
<td>greatest opportunity</td>
</tr>
<tr>
<td>Opportunity for building rapport</td>
<td>least</td>
<td>some</td>
<td>greatest</td>
</tr>
<tr>
<td>Potential negative respondent reaction</td>
<td>“junk mail”</td>
<td>“junk calls”</td>
<td>invasion of privacy</td>
</tr>
<tr>
<td>Interviewer control over interview environment</td>
<td>least</td>
<td>some in selection of time to call</td>
<td>greatest</td>
</tr>
<tr>
<td>Time lag between soliciting and receiving response</td>
<td>greatest</td>
<td>least</td>
<td>may be considerable if a large area is involved</td>
</tr>
<tr>
<td>Suitable types of questions</td>
<td>simple, mostly dichotomous (yes-no) and multiple-choice questions</td>
<td>some opportunity for open-ended questions, especially if interview is recorded</td>
<td>greatest opportunity for open-ended questions</td>
</tr>
<tr>
<td>Requirement for technical skills in conducting interview</td>
<td>least</td>
<td>medium</td>
<td>greatest</td>
</tr>
</tbody>
</table>