Web Survey Design

Mick P. Couper
(Survey Research Center, University of Michigan)

Session Objectives

- Raise awareness of importance of design for Web surveys or online data collection
- Illustrate design issues with empirical examples
- Offer selected guidelines to improve design of Web surveys
- Provide resources for further information on design
Selected Topics

- Input field choices (drop boxes, radio buttons, check boxes, text boxes, etc.)
- Grids or matrices
- Dealing with missing data
- Changing answers/back button
- Use of images
- Progress bars

Input Field Choices

- There are many different ways questions can be answered on the Web
  - E.g., radio buttons, check boxes, drop boxes, text fields, slider bars, ...
- The choice of input method and the design of the response options affect the answers obtained
- We illustrate this with some of our studies
  - Response format experiment
  - Input field length experiment
Response Format Experiment


- Members of opt-in panel randomly assigned to treatment in 2x3 design
  - Order of items:
    - One condition reversed order of response options
  - Format of question:
    - Radio button
    - Drop box, no responses initially visible
    - Drop box, 5 responses initially visible

Radio Button Version A

Which of the following nutrients is most important to you when selecting breakfast cereal? (Please select one)

- Protein
- Carbohydrates
- Sugar
- Fat
- Fiber
- Vitamin A
- Vitamin C
- Calcium
- Iron
- Vitamin E
- None of the above
Radio Button Version B

Which of the following nutrients is most important to you when selecting breakfast cereal? (Please select one)

- Vitamin E
- Iron
- Calcium
- Vitamin C
- Vitamin A
- Fiber
- Fat
- Sugar
- Carbohydrates
- Protein
- None of the above

Drop Box, None Visible: Before

Which of the following nutrients is most important to you when selecting breakfast cereal? (Please select one)
Drop Box, None Visible: After

Which of the following nutrients is most important to you when selecting breakfast cereal? (Please select one)

- Protein
- Carbohydrates
- Sugar
- Fat
- Fiber
- Vitamin A
- Vitamin C
- Calcium
- Iron
- Vitamin E

Questions about this survey?
Email us at unifile@umich.edu
or call toll-free 1.800.674.3375

Drop Box, Five Visible, Versions A and B
Implications

- We find clear primacy effects, consistent with the literature
- We find even stronger effects for visible options
- Visibility principle: What they see is what we get
- We have explored this further using eye-tracking equipment in a lab (Galesic et al., in press)
Percent of Time Fixated on Top Half of Response Options

Question on Child Qualities: Evidence of Primacy

Which one quality listed below would you say is the most desirable for a child to have?

- That he is interested in how and why things happen
- That he is considerate of others
- That he is responsible
- That he obeys his parents well
- That he gets along well with other children
- That he acts like a boy or she acts like a girl
- That he has self-control
- That he has good sense and sound judgement
- That he is neat and clean
- That he is honest
- That he tried hard to succeed
- That he has good manners

Based on average fixation durations across all respondents (the warmer the color, the longer the fixations), n=51
Different Response Styles

- We observed several different response styles
  - Considering all options and choosing the best answer:
  - Selecting the first option, then going through the list and updating the response:
  - Reading only part of the list, then selecting the answer:

- Overall, we find that respondents spend more time on the earlier options than on later options

Input Field Length and Format

- The size of a text/numeric input field conveys information to respondents

- Illustrative experiments
  - Experiment 1: Size of input field for narrative responses
  - Experiment 2: size of input field for numeric responses
  - Experiment 3: Use of masks/templates for currency amounts
  - Experiment 4: Date of birth
Experiment 1: Narrative Questions

- Dennis, deRouvray, and Couper (AAPOR, 2000)
- Used Knowledge Networks panel
- Panel members randomly assigned to large (n=164) or small (n=172) text boxes for three open-ended items:
  - Q0: Your Web TV can be used to surf the Web, chat with other viewers, program your VCR, email your friends and family, and more. How do you use your Web TV?
  - Q4: In your opinion, what are the principal benefits, if any, of ordering goods on the Internet?
  - Q10: In your opinion, what role, if any, should the U.S. government play in overseeing Internet security and privacy?

Results of Entry Box Experiment

All differences statistically significant (p<.05)

Source: Dennis, deRouvray, and Couper (2000)
Experiment 2: Input Field Size

- We hypothesized that the longer input field would encourage more “ill-formed” responses:
  - Answers in words (e.g., “twice”)
  - Ranges (e.g., “4-5”)
  - Estimates (e.g., “about 5”)

### Ill-Formed Entries by Field Length and Question

<table>
<thead>
<tr>
<th>Percent ill-formed</th>
<th>Doctor</th>
<th>Dentist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short field</td>
<td>0.9</td>
<td>1.2</td>
</tr>
<tr>
<td>Long field</td>
<td>4.1</td>
<td>3.3</td>
</tr>
</tbody>
</table>

Source: Couper, Kennedy, Tourangeau, and Conrad (in preparation)
Experiment 3: Use of Masks or Templates

- We hypothesized that the template would reduce ill-formed entries

<table>
<thead>
<tr>
<th>Percent of Ill-Formed Entries by Size of Input Field and Use of Mask</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Short field</strong></td>
</tr>
<tr>
<td><strong>No mask</strong></td>
</tr>
<tr>
<td>17.0</td>
</tr>
<tr>
<td>23.3</td>
</tr>
</tbody>
</table>

Main effects of field size (p<.01) and mask (p<.001) significant; interaction not significant

Source: Couper, Kennedy, Tourangeau, and Conrad, in preparation
Use of Masks Reduces Number of Ill-Formed Entries

Experiment 4: Entry of Dates

What is your date of birth?

MM/DD/YYYY

Short box

What is your date of birth?

MM/DD/YYYY

Long box

What is your date of birth?

MM/DD/YYYY

3 boxes

What is your date of birth?

Month Day Year

Drop boxes

Source: Couper, Kennedy, Tourangeau, and Conrad, in preparation
Experiment 4: Ill-Formed, Incomplete, or Missing Responses by Input Type

χ²=41.27, d.f.=3, p=<.0001

Experiment 4: Response Time by Input Type

Drop boxes significantly different from text boxes, t=8.47, d.f.=2,415, p<.0001
Implication of Input Field Experiments

- Size of field matters
  - If you want narrative responses, provide sufficient space
  - If you want numeric responses, match field size to expected input
- “Ill-formed” doesn’t mean wrong – it means more effort to code the data, and makes automatic branching or edit-checking difficult
- Masks and other visual cues help convey the format of the information desired

Grids or Matrices

- Grid or matrix questions – designed using HTML tables – are common designs for multiple items sharing the same response options
- Advantages:
  - Quicker response times
  - Higher inter-item correlations
- Disadvantages
  - Higher item missing data and breakoffs
  - More non-differentiation
  - More correlated measurement error
### Grid Example

This section is about computing tasks.

1. How often do you use a computer to do the following tasks? (Please select one response per row.)

<table>
<thead>
<tr>
<th>Task</th>
<th>Never</th>
<th>Once a year</th>
<th>Several times a year</th>
<th>Once a month</th>
<th>Several times a week</th>
<th>Every day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Write papers and reports</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do research using the web</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access online library resources</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use the web</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use email to communicate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buy, sell or manage personal stock</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participate in online work or projects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use email to communicate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Download music, pictures, videos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Play computer games</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please list here and describe below)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Additional comments

1. 1. Answer the following questions using a Likert scale.

   **Category 1**:
   23. Das interne Thema verbreitet, dass man in der Gesprächspartner
   seiner Meinung ändern kann.
   24. Verhaltensweisen erlernen
   27. Spannungsfelder und Widersprüche zwischen Teilnehmern
   einer Gruppe.
   29. Leute mit ähnlichen natürlichen Verhaltensmuster.
   30. Einzelne Menschen können besser
   31. Freunde finden
   32. Leute, die ich interessant finde, einfach ausfindig machen
   33. eine(n) zufriedenstellende Person(s) finden
   34. Leute treffen, mit denen ich mir im richtigen Leben mein Freude feierte.
   35. Menschen, die sich diese ihrer inneren Einheit, besser
   36. Leute treffen, die der gleichen Ansicht haben wie ich
   37. Menschen aus anderen Städten und Landstrichen treffen
   38. auch so geben, wie ich wirklich bin.
   40. Identitäre Absichten, die sich im richtigen Leben nicht haben lassen
   41. persönliche Details aus dem Privatleben anderer erfahren
   42. nur kleine Gedanken machen, wie ich gerade ausführe
   43. nicht mehr aus dem Haus gehen müssen.
   44. andere Leute mit ihrer Art erziehen

---

*UNIVERSITY OF MICHIGAN*
Research on Grids

- Our early work focused on the effect of grids on measurement
  - Example experiment 1
- Our current work is focusing on how to improve the design of grids to reduce the negative effects
  - Specifically, can the dynamic nature of the Web be used to guide respondents through grids?
  - Example experiment 2

Grid Experiment 1

- Series of 8 items on diet
- Three versions:
  - All 8 items on 1 page
  - Items split into 2 pages with 4 items on each
  - Each item on separate page, with 8 pages
- Cronbach’s alpha coefficients by format
Part-Whole Correlations for Reverse-Worded Items, by Format

- Negative correlations increase when items on separate pages
- Higher nondifferentiation (less differentiation among items) in grids
- Respondents less likely to notice reverse wording in grids

Grid Experiment 2:

- Galesic et al. (2007)
- Tested 3 navigation aids:
  - No change
  - Change of font color to light grey
  - Change of row background to light grey
- Tested in 3 separate grids
  - 8 items
  - 8 items
  - 18 items
Grid Navigation Examples

During the last week, how much of the following ingredients did you consume?

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>Much less than I should</th>
<th>Somewhat less than I should</th>
<th>About as much as I should</th>
<th>Somewhat more than I should</th>
<th>Much more than I should</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calcium</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Iron</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Sugar</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Vitamin C</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Caffeine</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Protein</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Fiber</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Percent dropping out

Differences not significant, $\chi^2=3.75, p=.15$

Source: Galesic et al. (2007)
Implications of Grid Experiments

- Grids are complex – require careful design
- While grids are popular among designers, they don’t seem to work as well for respondents
- One reason may be poor design rather than grids per se
- Use judiciously

Source: Galesic et al. (2007)
Dealing With Item Missing Data

- Depends on how the Web survey is designed
- Scrolling Web survey designs (all questions on one page):
  - Acts like paper – cannot alert respondent to missing data at each question (unless using client-side scripts)
  - Server-side script can alert respondent about missing items after submission
- Paging (interactive) Web survey designs
  - Various design strategies possible

Approaches to Missing Data

- General finding that providing an explicit “don’t know” or “refused” option increases the use of that option
- Forcing a respondent to answer every question is likely to produce breakoffs or poor data, and raises ethical concerns
- One solution is to provide an “opt-out” option, but make it less conspicuous than other choices
- Another is to provide pop-up reminder and motivator if respondent skips item, but don’t force an answer
Experiment on Reducing Missing Data

- DeRouvray and Couper (2002)
- Tested four variations of a “Decline to answer” option
  - “Decline” option
  - “Decline” option with prompt
  - “Decline” option in smaller, lighter font
  - No decline option, with prompt
- Sample size about 160 cases per cell
- Examined average item missing rates (decline+skip) for 12 items seen by all respondents

Example screens

Versions Tested and Missing Data Rates

- 1. “Decline to answer” option
- 2. “Decline to answer” option and reminder pop-up if respondent skips
- 3. “Decline to answer” option in smaller font and light gray color
- 4. No “decline to answer” option but reminder pop-up if respondent skips

Source: deRouvray and Couper (2002)
Missing Data: Implications

- Use the interactive nature of the Web to encourage respondents to answer all questions
- Do not force respondents to answer questions in order to continue
- Minimize the effect of explicit “don’t know” or “decline to answer” options
- Missing data on some items is better than complete or partial nonresponse (breakoff)

Changing Answers/Back Button

- Some surveys do not let respondents go back and change answers
  - Sometimes technical arguments are used
  - Sometimes data integrity arguments are used
- Respondents rarely go back and change answers
- Providing them with the opportunity to do so can increase the honesty of reporting
  - Also arguments for improved accuracy
  - But little research done on this
Should a “Back” Button Be Provided?

- Why not let respondent use back button on the browser to go to previous question or form?
- Some Web survey software systems have problems with browser back button, because of caching and unexpected commands, so provide their own back button
- Error generated if browser buttons used:

<table>
<thead>
<tr>
<th>Data Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>This document resulted from a POST operation and has expired from the cache. If you wish you can report the form data to recreate the document by pressing the 'reload' button.</td>
</tr>
</tbody>
</table>

Use of Back Button

During the past month, how often did you eat pasta?

You have used the browser buttons, please use the Next/Previous buttons below.

- Never
- Less than once a week
- Once a week
- Two or three times a week
- Four or five times a week
- Every day

[Last Screen] [Next Screen]
Back Button: Implications

- For ethical and data quality reasons, respondents should be able to change answers
- Provide respondents with a “back” button
- Use JavaScript to trap and redirect use of the browser back button
- Make the “back” button visually less prominent and out of normal flow of answering the survey
  - Permits but does not encourage use
  - Reduces accidental use

Use of Images

- It is easy to add images to Web sites and surveys
- Images are a key element on online tailoring materials
- Images believed to foster engagement
- However, images added to embellish or illustrate survey questions may have unintended effects on interpretation of the questions
- Effect of images on measurement error
  - Experiment 1: assimilation effects
  - Experiment 2: contrast effects
Image Experiment 1


- Respondents (n=2,400) from KN panel were independently randomized to one of four image conditions:
  - A high-frequency example
  - A low-frequency example
  - Both pictures
  - No picture

- This was done for each of a series of behaviors:
  - Shopping
  - Dining out
  - Overnight trips
  - Attending sporting events
  - Attending live music events
  - Listening to recorded music

Shopping: Clothing Store Version
Shopping: Grocery Store Version

How many times have you gone shopping since March 1st this year?
Enter an answer from 0 to 100

Knowledge Networks

Shopping: Results

- Significant (p < .01) effect on mean number of shopping trips in past month
- No interaction effects with respondent gender

Dining Out: Fancy Restaurant Version

About how many times have you eaten out since March 1st this year?
Enter an answer from 0 to 100

Dining Out: Fast Food Version

About how many times have you eaten out since March 1st this year?
Enter an answer from 0 to 100
Dining Out: Results

- Significant (p<.01) effect on mean number of times dined out in past month
- Similar significant effects for enjoyment of last meal and amount paid for last meal

![Bar chart showing mean number of trips for Dine-in and Fast food treatments.]


Image Experiment 2

- Effect of images on judgments of health
- Two picture conditions:
  - Fit woman jogging
  - Sick women in hospital bed
- Three location conditions:
  - In question area, above question
  - In header area
  - On prior screen, introducing the topic
- About 2,700 respondents from an access panel completed the survey
Healthy Image — In Question Area

Sick Image — In Question Area
Results

- Significant (p<.05) effect of images on self-reported health
- Percent of respondents reporting their health as excellent or very good
- Significant interaction of picture and location
  - Effect is reduced when picture in header
  - Suggests “banner blindness”

Eye-Tracking Study on Images

- Galesic, Tourangeau, Couper, and Conrad (in press)
- Goal: to further investigate the “banner blindness”, as well as contrast and assimilation effects
- Design:
  - Picture of a happy vs. sad woman
  - In the question area vs. in the header
  - Two questions (measured on 0-10 scale):
    - Evaluation of happiness
    - Evaluation of sadness

Source: Couper, Conrad, and Tourangeau (2007)
Looking at the Pictures, by Position: Happy-Sad

- 92% looked at the pictures in the question area, 70% in header ($\chi^2(1)=8.37, p<.01$)
- More time spent looking at the pictures in the question area:

![Bar chart showing gaze duration in msec for Happy and Sad pictures in header and question area.

Main effect of position: $F(1,105)=10.04, p<.01$ (for log-transformed data)]

Evaluations of Happiness, by Picture

- Contrast effect when respondents fixated on the picture, assimilation when they did not; similar results for sadness

![Graph showing evaluations of happiness for Sad and Happy women.

Interaction: $F(1,104)=8.29, p<.01$]

Note: “Looking” was defined as spending at least 100 msec on a picture (cf. Lohse and Johnson, 1996; Stark, 1994; Josephson and Holmes, 2002).
Implications of Images for Web Surveys

- People attend to images
- Consciously or unconsciously, they can affect answers, altering
  - The construal of the category,
  - The specific instances that come to mind,
  - The standard to which target item is compared
- Images are strong contextual cues, perhaps even stronger than prior questions
- We find no evidence that inclusion of images reduces breakoff
- Use images in surveys with caution – only when the image is germane to the question

Progress Bars

- Many believe that providing indicators of progress through the survey are both desired by respondents and useful in reducing breakoff
- Several empirical studies have demonstrated at best marginal positive effects and at times significant backfire effects
- On balance, progress indicators do not appear to help
- What is more important is what respondents are told up front about the length of the survey
General Conclusion

- It’s not just the words that matter in Web and other survey modes – the visual and interactive elements affect answers too
- Web surveys offer many more tools for the designer
- Care must be taken with design
- Consult a survey professional

For More Information

- **SELF-PROMOTION ALERT**: Read the book (available September 2008)
- Visit [www.websm.org](http://www.websm.org) for more on Web survey methodology
Thank You

Cited References