

Best Practices in Survey Design

Shelley Strowman, Pottruck Technology Resource Center, Simmons College

Why is it important to create a good survey?

- Keeps respondents interested
- Reduces respondent error and frustration
- Facilitates data entry and analysis
- Enables you to use the survey again with few or no changes needed
- Your data are only as good as your survey ("garbage in, garbage out")

Types of Questions

Closed vs. Open-Ended Questions

1. *Closed*: Respondents are asked to choose one or more responses from a list
 - a. Example: "In your opinion, which of these candidates is most trustworthy?"
 1. George Bush 2. John Kerry 3. Ralph Nader 4. Other 5. None
2. *Open*: Respondents write in their answers
 - a. Numeric open-ends (e.g., "What is your age? _____")
 - b. Text open-ends (e.g., "Other: _____" or "What do you like most about this candidate?")

Single vs. Multiple-Response Questions

1. *Single response*: Respondents must choose only one answer
2. *Multi-response*: Respondents are told to "choose all that apply"
 - a. Note that during data entry, response alternatives must be treated as a series of individual items coded as yes and no (i.e., respondent checked vs. did not check that box)

Quantitative vs. Categorical Items (Affects the Types of Analyses You Can Do)

1. *Quantitative*: Ratio or interval data
 - a. You can calculate a mean and standard deviation on these items
 - b. Example: "What is your age?" _____ years
2. *Categorical*: Ordinal or nominal data (numeric codes that stand for categories)
 - a. You cannot calculate a mean and standard deviation on these data
 - b. Example: "Which of the following categories best describes your age?"
 1. 18-25 years 2. 26-35 years 3. 36-45 years 4. 46-55 years 5. 56+ years

Tips for Designing a Good Survey

1. Organization/ Layout of the Survey

- a. *Make sure the survey is well-organized*
 - Group similar items into sections using subheadings if necessary
 - Order items logically within sections
 - Keep skip patterns to a minimum and use clear skip instructions
 - Skip instructions tell respondents to skip to another item depending on their response (e.g., "If you answered no, skip to Q8")
 - Put skip instructions in parentheses and consider using arrows
 - Put general instructions in bold
- b. *Good questionnaire flow keeps respondents interested and happy*
 - Appearance of the survey can make a huge difference
 - Keep the layout uncluttered
 - Number the questions (e.g., q1, q2a, q2b, etc.)
- c. *Put certain types of items in specific areas of the survey*
 - Demographic items are often placed last because these are personal and you don't want to offend anyone up front
 - May also want to put sensitive items last or at the end of the relevant section
- d. *Keep the survey to a reasonable length*
 - Efficiency is more important than length
 - Avoid items that are off topic or redundant
 - Avoid a long series of items about the same topic

2. Survey Items

- a. *Be specific, without getting too wordy or lengthy (try to keep to 20-25 words)*
 - Bad: "Did you vote in last election?"
 - Good: "Did you vote in the general election in November 2000?"
- b. *Avoid biased/ loaded questions*
 - Bad: "John Kerry is passionate about protection of our environment. Do you plan to vote for him or George Bush?"
 - Good: "Which one of the following presidential candidates are you planning to vote for on Nov 2"?

a. George Bush b. John Kerry c. Ralph Nader d. Other e. None
- c. *Avoid double-barreled items (items that ask about more than one topic)*
 - Bad: "Please rate your agreement or disagreement with this statement: The U.S. was justified in going to war in Afghanistan and Iraq."
 - Good: "Please rate your agreement or disagreement with this statement: The U.S. was justified in going to war in Iraq."
- d. *Avoid double negatives*
 - Bad: "I am not satisfied with the choice of candidates in this year's presidential election"

a. Strongly Disagree b. Disagree c. Neutral d. Agree e. Strongly Agree

 - Good: "I am satisfied with the choice of candidates in this year's presidential election"

a. Strongly Disagree b. Disagree c. Neutral d. Agree e. Strongly Agree

- e. *Use closed-ended items for sensitive topics*
- Bad: "What was your total household income, before taxes, in 2003? _____"
 - Good: "What was your total household income, before taxes, in 2003?"
 - a. < \$20K
 - b. \$21-40K
 - c. \$41-60K
 - d. \$61-80K
 - d. \$81-100K
 - e. \$101K+
- f. *Things to consider when items involve recall*
- Try to limit recall to the recent past, especially for minor events or infrequent behaviors
 - Be very clear with your definition of the time period
 - For items that require recall of a lot of information, provide choices to aid recall rather than leaving the item open-ended
- g. *Use of single response vs. multiple response items*
- Be very clear in your instructions about whether respondents should choose only response or all that apply
 - If using a single response item, make sure respondents will be able to choose only one response (and won't feel frustrated by the forced choice)
 - Remember that multi-response items must be treated as a series of separate items in the data file

3. Response Alternatives

- a. *Response alternatives should be exhaustive and mutually exclusive*
- Make sure your responsive choices allow for all possible answers (they are exhaustive)
 - Provide a category for "other" if necessary
 - Include a category for "none" when needed
 - Responses choices should also have no overlap (they are mutually exclusive)
- b. *Likert scale items (attitudinal questions, often involving disagree-agree scales)*
- Ideally, these should include 5 to 9 response choices
 - Provide a neutral middle point, unless you are using "forced-choice" items (in which case, you should instruct respondents to select the one "best" answer)
 - Label all anchor points so all respondents are applying the same meaning to the numbers
 - The scales should be balanced on both sides
 - It's best to assign numeric codes such that 1=least/worst/disagree, 5=most/best/agree
- c. *Use of "Don't Know" (and related options such as "Not Sure" or "Not Applicable")*
- Include a "Don't Know" option when it's important to measure this (e.g., on a knowledge survey) and when it's realistic that some respondents actually won't know
 - "Don't Know" should also be included if respondents might get frustrated without this option (otherwise, they'll probably just skip the item anyway)
 - On the other hand, "Don't Know" is similar to missing data in many cases - so you will lose some valid responses if you include this option
 - Think beforehand how you want to analyze "Don't Know" responses - will you assign these missing value codes (e.g., 9 or 99) and define them as missing values?

- d. *Avoid using "never" and "always" unless they are clearly possible answers*
- Bad: "In the past month, how often did you do canvassing for a campaign?"
a. Never b. Rarely c. Sometimes d. Often e. Always
 - Good: "In the past month, how many times did you do canvassing for a campaign?"
a. 0 times b. 1-2 times c. 3-4 times d. 5+ times
- e. *Avoid using just "yes" and "no" if you can be more specific*
- Bad: "Are you planning to vote on Nov 2?"
a. Yes b. No
 - Good: "How likely is it that you will vote on Nov 2?"
a. Very unlikely b. Somewhat unlikely c. Somewhat likely d. Very likely
- f. *Other pointers about response choices*
- Label anchor points to minimize subjectivity
 - Make sure the response choices match the question
 - Don't give too many choices or respondents might not read the whole list
 - When you want to know the frequency of behavior, it's generally better to define the frequency rather than use "often," "sometimes," etc.
 - For sensitive items (e.g., alcohol use), use response choices that will avoid making the respondent feel their response is extreme

4. Anonymity and Confidentiality

- a. *Keep the survey anonymous if you want to get the most honest responses*
- b. *If you can't offer anonymity, at least offer confidentiality*
(e.g., only a limited number of people will see their data, no identifying information will be revealed, only groups of answers will be reported, etc.)

6. Pre-Test the Survey

- a. *Administer the survey to a small group of respondents*
- Include some people who are familiar with the topic and some who are not
- b. *Ask for feedback in terms of:*
- clarity of instructions, items, and response choices
 - lack of bias in items
 - response choices being exhaustive and mutually exclusive
 - skip patterns – were there too many, was it clear how to skip?
 - ambiguous terms defined well
 - questionnaire flow
 - appearance of survey (readable vs. cluttered, professional-looking?)
 - length of items
 - length of survey
 - sensitive items asked appropriately
- c. *Look for potential problems including:*
- Respondents aren't following instructions → make the instructions clearer
 - Too many "Don't Know" responses → reword those items or remove "Don't Know" options to force a response
 - Too many items left blank → reword or delete those items, make sure response choices are exhaustive and mutually exclusive

- Respondents choose more than one response for single response items → make sure instructions are clear about choosing one vs. multiple answers
- Surveys aren't finished → make survey shorter
- Items have little variance → reword item or provide more response alternatives
- Respondents aren't following skip patterns correctly → make skip instructions clearer or reword/reorder items so skip patterns are less complicated
- Several respondents provide the same answer to "Other: _____" → add that response to the list of response alternatives

6. Other Suggestions

- a. Pre-code the survey* – this will help you identify problems and will make data entry easier
- b. Include a cover letter* that explains why it's important to respond and how the survey results could impact their lives
- c. Send reminders and use incentives* to increase response rate – low response rates can be a problem if respondents differ from non-respondents

References

- DeVellis, R.
(1991). *Scale development: Theory and applications*. Newbury Park, CA: Sage.
- Fowler, F.
(1993). *Survey research methods*. Newbury Park, CA: Sage.
- Mangione, T.
(1995). *Mail surveys: Improving the quality*. Newbury Park, CA: Sage.
- Patten, M.
(2001). *Questionnaire research: A practical guide*. Los Angeles: Pyrczak Publishing.
- SPSS Inc.
(1998). *Survey research using SPSS*. Chicago: SPSS Inc.