

### Some Thoughts on Writing Questions

Introductory Note: My experience with trying to teach the writing of questions has been that the big picture of what we are trying to achieve with survey questions often gets missed. Sometimes people get lost in one small issue (e.g. Which is more valid five or seven point scales?) and may miss the bigger question of whether they should even be using such scales, and what the wording should be. The biggest difficulty I have faced in writing questions is being able to go back and forth from the information in front of me that some things should be fixed, and to what aspect of knowledge needs to be drawn upon to design a fix. One of the things that I have learned to do over my career is not to think about everything at once, but try to think in terms of layers of concern. Today's presentation tries to identify some of the multiple layers one needs to think about in a practical way. Fowler and Conseza provide good background as well as examples and my book also provides background for where I hope to do today.

- A. Our objective is to produce an inquiry that will allow us to produce the distribution of a characteristic in the population.

Think about constructing a "variable." Otherwise, why ask the question?

e.g., What percent of the U.S. adult population is unemployed?

- B. Error concerns: mostly measurement, but also non-response.

e.g., "Have you every shoplifted something?" vs. "Have you ever taken anything from a store without paying for it?"

- C. Questions I try to ask myself when writing survey questions and working with clients who ask me for help with their proposed questions. (For this to make sense it helps to think about trying to help someone else with their questions. I use this as a mental starting point to understand some of the general constraints that need to be dealt with.

1. Does the question you have proposed require an answer?
2. To what extent do survey recipients already have an accurate ready-made answer for the question they are being asked to report?
3. Can people accurately recall and report past behaviors?
4. Is the respondent willing to reveal the requested information?
5. Will the respondent feel motivated to answer each question?
6. Is survey information being collected by more than one mode?
7. Is changing a question acceptable to the survey sponsor?

D. I find it useful to think of questions as asking for one of four types of information, but over the years have used it somewhat less than in the past.

1. Attitudes—What one wants or prefers.
2. Beliefs—What one thinks to be true.
3. Behavior—What one does or has done.
4. Attributes—What one is.

Generally, answers to attitude and belief questions, especially abstract ones, exhibit more unintentional measurement error than do answers to attribute and behavior questions. But, recalling behavior is sometimes tricky—we'll talk about this later.

E. Fundamentally, there are four types of question structures.

1. *Open-ended*: No answer choices are offered. Frequently, it is difficult to create variables because of the vague or imprecise stimulus.
2. *Close-ended with ordered response categories (ordinal scale)*: Answer choices are offered and form a continuum which a respondent must place her/himself on.
3. *Close-ended with unordered response categories (nominal scale)*: Answer choices are offered, but do not form a continuum of choices.
4. *Partially close-ended*: Respondents have answer choices but, in addition, can offer their own open-ended answer (usually associated with close-ended unordered).

Are there other types you can think of?

F. Will people provide honest answers?

1. Not necessarily and sometimes respondents aren't really aware that their answers aren't entirely truthful. There are tools for decreasing likelihood of untruthful answers—context, embedding, and other ideas associated with how conversations occur. Sudman and Bradburn have written good books on this topic.
2. There are limitations to what can be done to achieve truthfulness.

G. Layers of complexity: the question building process.

1. *First level of complexity*. Thinking about wording in two locations—the query and the response choices.
2. *Second Level of Complexity*. Ordering questions and response categories.

3. *Third level of Complexity.* Mode problems and human tendencies.

H. First Level Problem—People don't draw information from just the query.

- 1) Words in query
- 2) Words in the categories
- 3) Number of categories
- 4) Position of categories
- 5) An experimental examples: estimates of time spent by students studying, on a computer and watching television.

I. Meaning from words (see text for elaboration on each).

1. Vaguely worded questions and responses.
2. Abbreviations or jargon.
3. Too much precision.
4. Bias from slanted introduction.
5. Bias from unequal comparison.
6. Bias from unbalanced response choices.
7. Bias from tone of the question.
8. Objection to providing information.
9. Objectionable statements.
10. Questions that are too difficult.
11. Doubled-barreled questions.
12. Answers are not mutually exclusive.
13. Assumes too much knowledge.
14. Inaccurate statements.
15. Inappropriate time references.
16. Others.....the list goes on.

J.. Second level wording problems—ordering questions and categories.

1. A questionnaire is a conversation—the meaning depends upon what has been said before, and mutual understandings. (cognitive concerns). Grice argues that four principles structure the exchange of information between people in a conversation. (Norbert Schwarz provides elaboration).
  - a. speakers should not say things they believe to be false (truthful).
  - b. speakers should make comments that are relevant to the purposes of the conversation (relevant).
  - c. speakers should make their contributions as informative as possible (informative).
  - d. speakers should express themselves as clear as possible (clarity)
  - e. See Thinking about answers by Sudman, Bradburn and Schwarz, Jossey-Bass, 1996).
2. Research results on multiple topics that illustrate these concepts.
  - a. Addition/subtraction effects (Schuman and Presser, 1981)  
abortion questions--subtractions  
how is life questions—carryover

- b. Primacy vs. recency  
Kosnick/Alwin  
Dillman et al. 84 experiments.
  - c. Anchoring on the first item  
Cheating at WSU vs. across the United States.
  - d. Norm-=of-even-handedness.  
Communist reporting, restricting trade, etc.
3. How do we diagnose problems in questions related to comprehension and process to identify problems in question wording? A four-step cognitive model of how people answer questions. (See Tourangeau, Roger, et al., ***The Psychology of Survey Response***. Cambridge University Press, 2000.)

- a. *Comprehension*: What am I being asked to do?

(Cognitive)

- Where is the question on this page?
- Where are the answer choices?
- In what order do I read this information in front of me?
- Carry-over and back five effects.

(Motivational)

- What's the question; how hard is it to figure that out?
- Is a question actually stated in plain English?
- How many parts does it have?
- Do I have to go somewhere else to figure out exactly what is being asked?
- How hard is it to find that information?
- Is a question even being stated?

Note: Process for self-administered questionnaires involves more than that described by Tourangeau and Rasinski for interview surveys: must understand the layout as well as the substance of the questions.

- b. *Retrieval*

Okay, now I understand the question.  
Can I provide an answer from memory?  
Do I need to check some records? Where are they?  
Easier when prior questions required retrieval of information on same topic.

- c. *Judging*

Okay, this is a satisfactory answer; it's close enough.  
(Or) No, we've got to have this exact ...

Do units of measurement have to be converted?  
*Have I already answered this question? (Addition and subtraction effects.)*  
Where do I anchor on a scale?

d. *Reporting*

Where do I provide the answer on this page?  
Is it clear where an answer is supposed to go?

e. To complete these four cognitive steps adequately, the respondent must be **motivated**. For example:

How hard am I willing to work to comprehend the question?  
How much am I willing to do to retrieve the needed information?  
How hard am I willing to work to make a judgment?  
How much effort am I willing to make to report (write) my answer?

...which may be influenced by such things as:

Do I want to help the sponsor?  
What did the cover letter say?  
Does this letter writer have an attitude?  
Is the purpose of this study worthwhile?  
Do I even know how these data get used?  
Will results of this study come back to "haunt" me?  
Is this a competently constructed questionnaire?  
Is it harder to understand than it needs to be?

An important question: *Do coercive letters, questionnaires that are really hard to comprehend, and instruction booklets filled with information that most respondents define as useless, result in people minimizing their efforts to comprehend, retrieve, decide and report?*

f. What survey design is all about: We must **motivate** people to adequately complete the **cognitive** steps necessary for answering questions accurately and returning the questionnaire.

i. Normative issues:

(1) *Social desirability*: The tendency for people to provide answers that put themselves in a good light with the person who asks the questions—a normal conversational behavior. (health status, drinking alcohol)

(2) *Acquiescence*: A cultural characteristic; in most cultures it is easier to agree with someone than it is to disagree.

(3) Some practical issues affected by cognitive, normative considerations:

- ◆ Use of agree/disagree scales.
- ◆ Issues on which people do not have well-informed opinions.
- ◆ Focusing only on expressed attitudes and opinions vs. behaviors.
- ◆ Cognitive design of questions.
- ◆ Whether to add Don't know category at end or in middle of scale.

(4) The special problems of scales:

- ◆ 3, 5, 7, 9, 10 or ? categories.
- ◆ Label or don't label?
- ◆ Sources of influence:
  - Signals from labels
  - Signals from category location
  - Signals from grouping as a series
  - Signals from spacing

K. Principles from Chapter 2 (***Mail and Internet Surveys***) that combine wording, order, and other structural issues together. These will not be covered individually in the short course; they are here for convenient reference (the book does not provide a table of contents for the principles).

Principle 2.1: Choose Simple over Specialized Words

Principle 2.2: Choose as Few Words as Possible to Pose the Question

Principle 2.3: Use Complete Sentences to Ask Questions

Principle 2.4: Avoid Vague Quantifiers When More Precise Estimates Can Be Obtained

Principle 2.5: Avoid Specificity That Exceeds the Respondent's Potential for Having an Accurate Ready-Made Answer

Principle 2.6: Use Equal Numbers of Positive and Negative Categories for Scalar Questions

Principle 2.7: Distinguish Undecided from Neutral by Placement at End of Scale

Principle 2.8: Avoid Bias from Unequal Comparisons

Principle 2.9: State Both Sides of Attitude Scales in the Question Stems

Principle 2.10: Eliminate Check-all-that-apply Question Formats to Reduce Primacy Effects

Principle 2.11: Develop Response Categories Which Are Mutually Exclusive

Principle 2.12: Use Cognitive Design Techniques to Improve Recall

Principle 2.13: Provide Appropriate Time Referents

Principle 2.14: Be Sure That Each Question Is Technically Accurate

Principle 2.15: Choose Question Wordings That Allow Essential Comparisons to Be Made with Previously Collected Data

Principle 2.16: Avoid Asking Respondents to Say Yes in Order to Mean No

Principle 2.17: Avoid Double-Barreled Questions

Principle 2.18: Soften the Impact of Potentially Objectionable Questions

Principle 2.19: Avoid Asking Respondents to Make Unnecessary Calculations

L. Third Level Wording Problems—differences among survey modes.

1. Fundamental Differences—interview vs. self-administered surveys (see Chapter 6)

- ◆ Presence vs. absence of interview
- ◆ Visual vs. Aural communication.
- ◆ Pace and control of answering process.

2. Mechanisms that may produce different answers

Normative:

- ◆ Social desirability
- ◆ Acquiescence
- ◆ Norm of even handedness (question order)

Cognitive—time pressures, control of sequence:

- ◆ Addition and subtraction
- ◆ Primacy vs. recency
- ◆ Top-of-the-head responses
- ◆ Memory: short vs. long term

3. Surveyors ask questions in different ways.

4. Types of mode effects documented in the literature

- a. Social desirability in interviews.  
How is your health  
How often do you drive car after drinking alcoholic beverages.
- b. Acquiescence in interviews.  
Agree-disagree  
How to get seat-belt laws to work
- c. Normative evenhandedness greater in interviews.
- d. Non-normative, addition and subtraction effects greater in interviews
- e. Extremeness on the telephone (See book for considerable detail on this effect.
- f.. Primacy vs. recency, picking the first offered categories (primacy) because all answers were compared to it on mail vs. picking last offered answer (recency) on telephone because first ones less memorable.
- g. Impacts of visual layout b(we'll cover this later.