



The following slides are the property of their authors and are provided on this website as a public service. Please do not copy or redistribute these slides without the written permission of all of the listed authors.

Session 4

Question Answering: Best Practices for Designing Closed-Ended Responses

Dr. Nicholas A. Valentino
University of Michigan



معهد البحوث الاجتماعية والاقتصادية المسحية
Social & Economic Survey Research Institute

Outline for this session

- I. Surveys are conversations
- II. How MANY response options should we offer?
- III. WHICH options should we offer?
- IV. Should we offer a MIDDLE category response?
- V. Should we offer a DON'T KNOW response?
- VI. Response ORDER effects

I. Remember! The survey is a conversation

- Answering questions is a cognitive process.
 - 1. Interpret the question, define terms
 - Most or all respondents must agree on the definition!
 - 2. Search memory for relevant information
 - 3. Combine many considerations into a “summary” judgment
 - 4. Translate judgment into the options provided

Choosing the correct response options can reduce SATISFICING

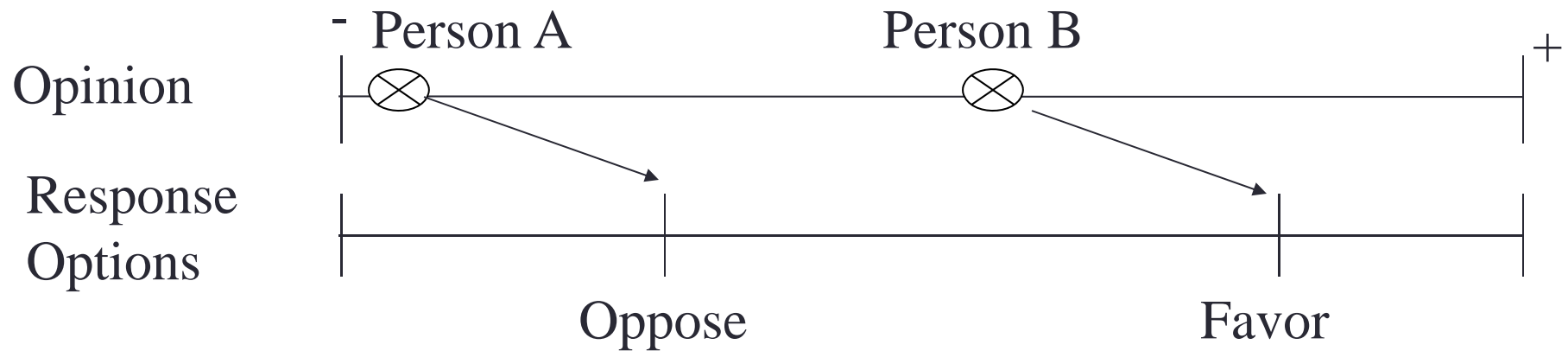
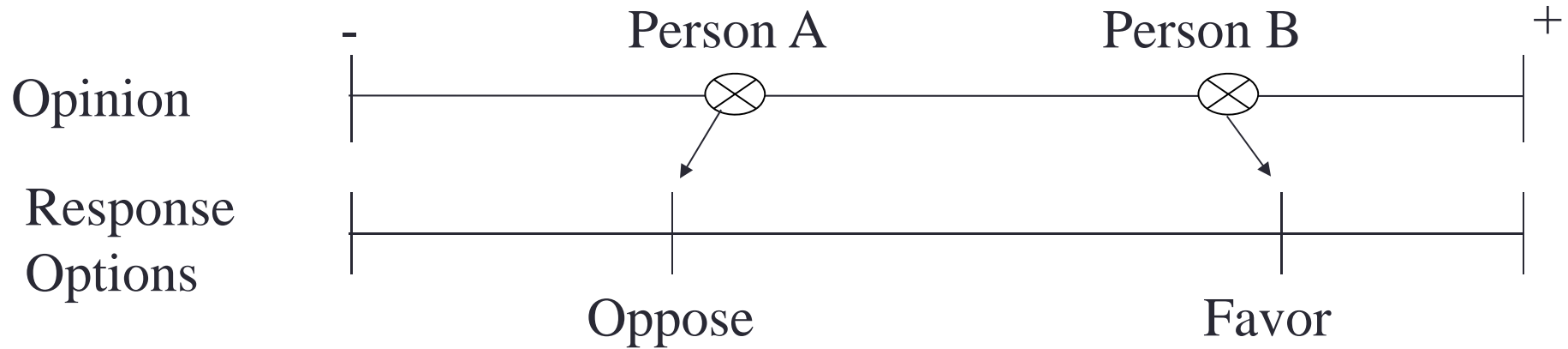
- Selecting the first reasonable response
- Being “agreeable” (Dr. Hutchings covered this)
- Saying “don’t know”
- Choosing options randomly



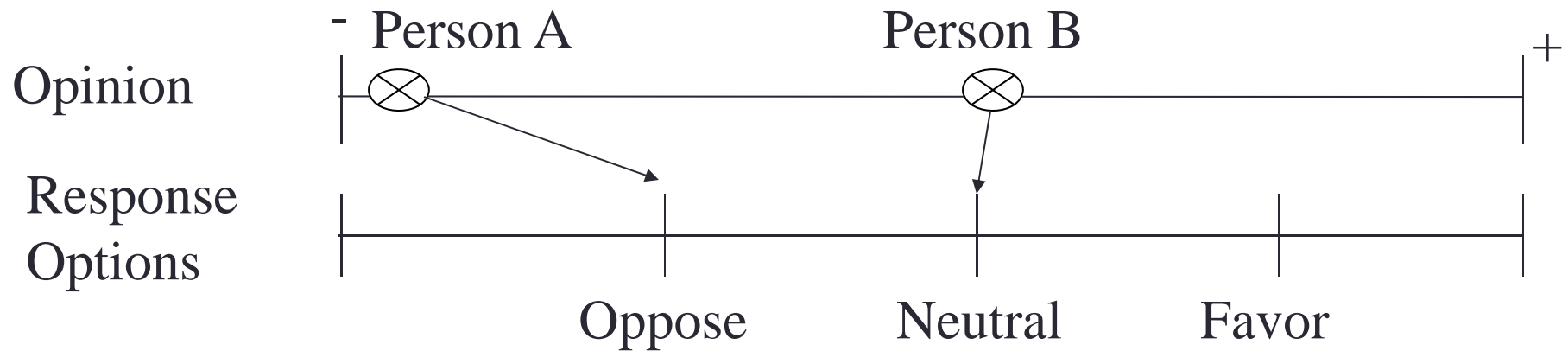
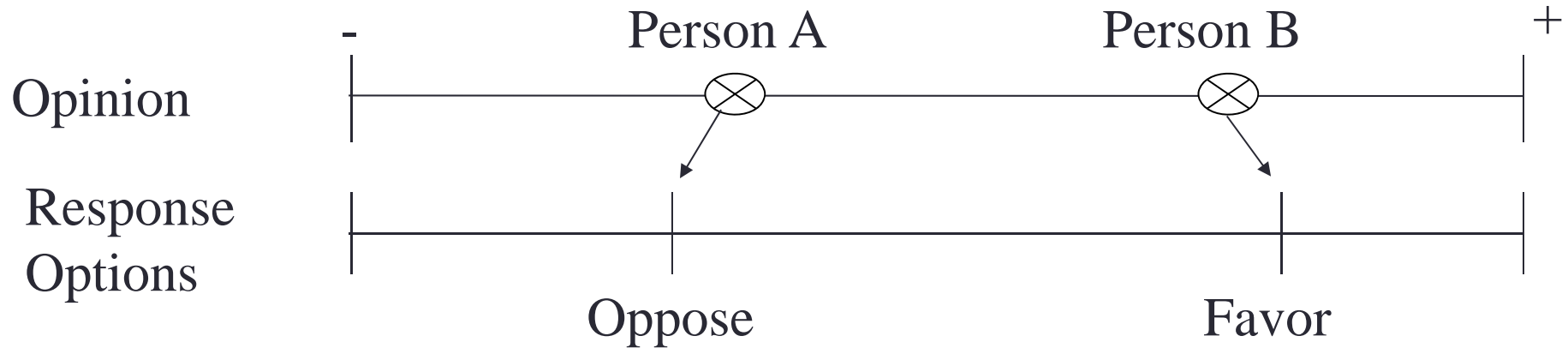
II. So how MANY response choices should we offer?

- Enough to allow respondents to translate their actual opinions precisely into an offered choice

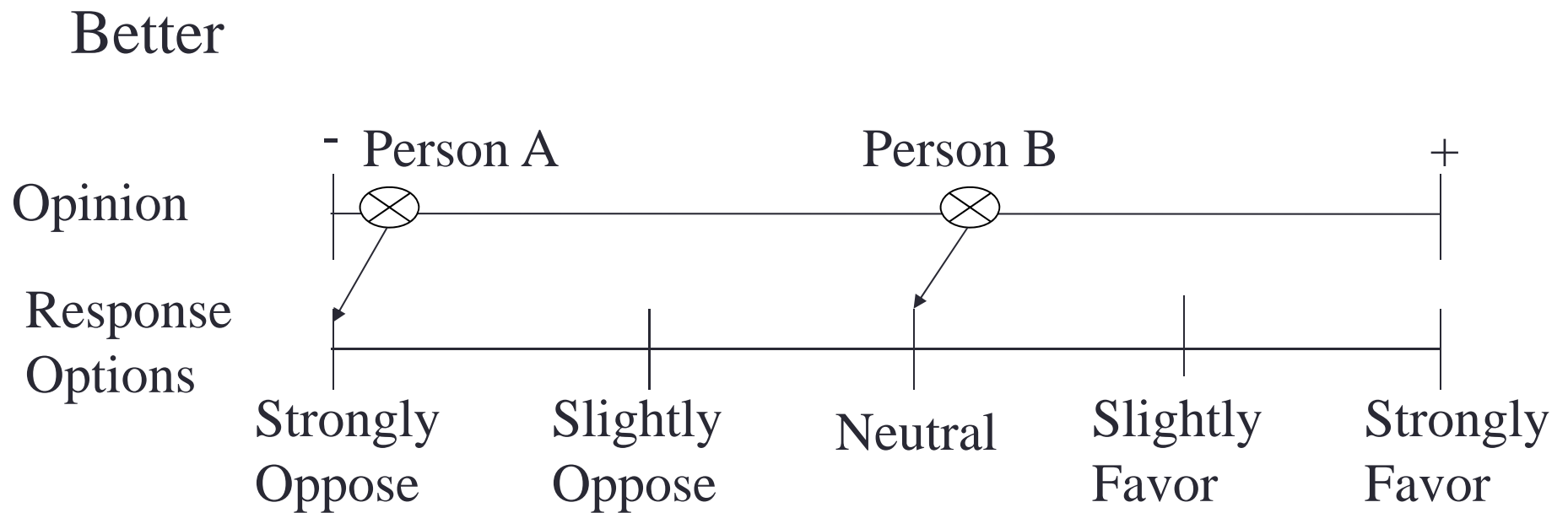
Translating opinions into response options: “Do you favor or oppose instruction in English in all schools in Qatar?”



Translating opinions into response options: “Do you favor or oppose instruction in English in all schools in Qatar?”

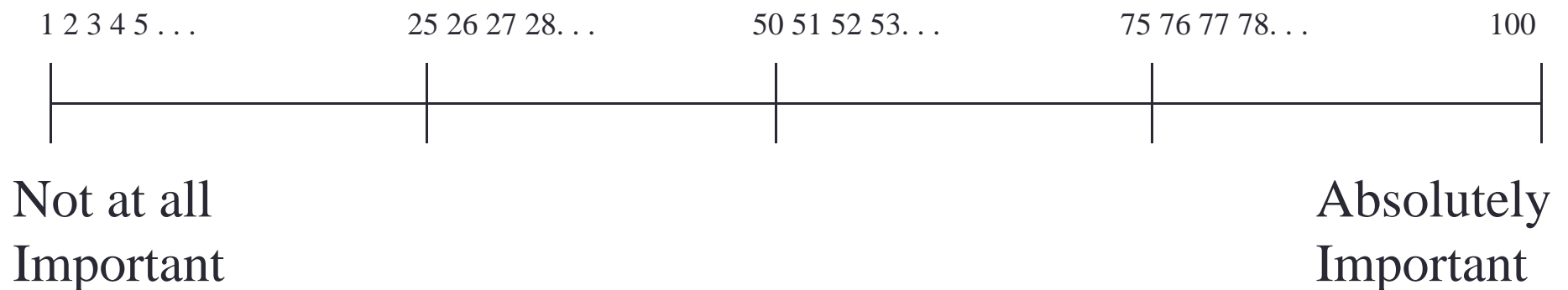


Translating opinions into response options: “Do you favor or oppose instruction in English in all schools in Qatar?”



Too many?

How important is it for you to live in a country that is governed democratically? On this scale where 1 means it is “not at all important” and 100 means “absolutely important” what position would you choose?



Conclusions about the ideal NUMBER of options

- 1) Mapping: More is better
- 2) Information gain: More is better
- 3) Clarity of meaning: Too many become ambiguous**
- 4) Time: More options- takes longer to answer question**

Conclusion: Increasing precision up to a certain number (5 or 7 options), decreasing precision thereafter.

It is best to use 5 choices when scale is UNIPOLAR

- Extremely
- Very
- Moderately
- Slightly
- Not at all

- Definitely will
- Probably will
- Might or might not
- Probably won't
- Definitely won't

- A great deal
- A lot
- A moderate amount
- A little
- None at all

- Always
- Most of the time
- About half the time
- Sometimes
- Never

Use 7 when scale is BIPOLAR

- Extremely good
- Moderately good
- Slightly good
- Neither good nor bad
- Slightly bad
- Moderately bad
- Extremely bad

If you insist on 5:

- Like a great deal
- Like a moderate amount
- Like a little
- Neither like nor dislike
- Dislike a little
- Dislike a moderate amount
- Dislike a great deal

- Excellent
- Good
- Fair
- Poor
- Very poor

III. WHICH response options?

- Labels must be EXHAUSTIVE & MUTUALLY EXCLUSIVE
- Labels must be easy for respondent to interpret
- Good labeling makes “translating” true opinion into response options simpler.

III. Choosing the right response options

- Options must be EXHAUSTIVE:
 - Must cover entire range of possible responses.
 - “What is the most important issue facing Qatar today?”
 - 1. Energy policy
 - 2. Environment
 - 3. Jobs
 - 4. International Security
 - 5. Education
 - 6. Roads and infrastructure

III. Choosing the right response options

- Labels must be MUTUALLY EXCLUSIVE
 - Scale points should be ordered and not overlapping.
- How strongly do you feel that only educated people should be involved with government.
 - Very strongly
 - Strongly
 - Not too strongly
 - Not at all

III. Choosing the right response options

- Points on the scale must have STABLE MEANING.
 - “Now I will ask you to use a “feeling thermometer” that runs from 1-100 to evaluate various groups, people, and places. A score of 1 means you feel very cold toward the person, group, or place, while a score of 100 means you feel very warmly about the person, group or place.” On this 1-100 point Feeling Thermometer, how warmly do you feel about **Doha**?”
 - Answers are affected by what *season* we are in.
 - People will gravitate toward multiples of 5 (5, 10, 25, 50, 75, 100)

Help Participant Translate Opinions into Response Options

“I received good medical care at the hospital.”

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
----------------	-------	---------	----------	-------------------

Step 1: How good was the medical care I received?

Excellent	Good	Fair	Poor	Very Poor
-----------	------	------	------	-----------

Step 2: Map answer onto response choices

Very Poor = Strongly Disagree


Poor = Disagree or strongly disagree

Fair = ? (not neutral)

Good = Agree? Strongly agree? (because I’m certain it was “good”)

Excellent = Strongly agree (but “good” doesn’t seem to capture all my enthusiasm)

= Strongly Disagree? (it wasn’t just good; it was excellent!)

- 
- 1) The “translation” process can be difficult for respondents and yields imprecise reflections of the underlying opinion.

Conclusion: Use response options that **match** the construct in the question. This is another reason to generally avoid “agree-disagree” scales.



So a better way to ask is. . .

- “Was the medical care you received at the hospital excellent, good, fair, poor, or very poor?”

Think about the *meaning* of response options

- Do you find it “acceptable” versus “do you support” it?
- “After the jury has decided a lawsuit, having the judge– instead of the jury– set the amount of damages awarded.”
 - “Very or somewhat acceptable” 60%
 - “Strongly or somewhat support” 17%

IV. Should we offer a a middle category?

- Benefit: Some people actually feel ambivalent
- Cost: Some might use this option to “satisfice”
- >20% will use middle category if offered
- Makes less difference to people who feel strongly.
- If intensity is important, leave it in.
- Most research suggests it does NOT undermine validity to include a middle category.

IV. Should you give a middle category? (Yes, most of the time!)

- In general, do you think public opinion polls are a good thing for the country or a bad thing? (GALLUP)
- Good thing 87%
- Bad thing 8



- **In general, do you think public opinion polls are a good thing for the country or a bad thing – or don't they make any difference one way or another? (UM)**

- **Good thing** **39%**
- **No difference** **46**
- **Bad thing** **10**

V. Offer A Don't Know option?

- Argument in favor: Reduces expression of “non-attitudes”
- Argument against: Can discourage true opinion expression.

Avoiding “non-attitudes”

- Washington Post poll: Support for the Public Affairs Act in 1995.
- “Some people say that the Public Affairs Act should be repealed. Do you agree or disagree with this idea?”
 - Repeal 24%
 - Do not repeal 19%

But “Don’t Know” could be due to social desirability pressures

- “How often have you used illegal drugs in your lifetime?”
- 1. Never
- 2. Occasionally
- 3. Often
- 4. Frequently
- 5. Don’t Know

V. Offer A Don't Know option?

- Research suggests explicit Don't Knows do more harm than good.
- People say “Don't Know” for many reasons OTHER than when they have no information.
- Studies show they do not improve predictions.
- Leave out explicit Don't Know UNLESS:
 - The issue is VERY new, most people have not thought about it.

VI. A Theory of Response Order Effects

- Visual Presentation – Primacy
 - Satisficing: take the first plausible choice
- Oral Presentation – Recency
 - Memory- The last one spoken is easiest to remember

VI. Order of response options

- Example of a **PRIMACY** effect
- Candidate preference affected by ballot *order*.
 - In 2000, Florida lists Bush first.



Bush v. Gore 2000

Gore:

57% in first place

54% in last place

Bush:

41% in first place

39% in last place

SOURCE: Pasek, J., Schneider, D., Krosnick, J. A., Tahk, A., & Ophir, E. (Under Review).
Prevalence and Moderators of the Candidate Name-Order Effect: Evidence from All
Statewide General Elections in California.

VI. Order of response options

- When response categories are presented orally, there is a **RECENCY EFFECT**
 - The last category spoken is chosen more often than when it is spoken first.

Recency Effect

(Schuman & Presser, 1981)

Some people say that we will still have plenty of oil 25 years from now. Others say that at the rate we are using up our oil, it will all be used up in about 15 years. Which of these ideas would you guess is most nearly right?

Percentage giving “plenty” response

	<u>“Plenty” first</u>	<u>“Plenty” last</u>	<u>X^2</u>	<u>p</u>
SRC-79 Jan.	63.5%	77.3%	13.00	<.001
	(293)	(273)		
SRC-79 April	60.7%	68.8%	4.17	<.05
	(443)	(218)		

When are Response Order Effects Largest?

When people lack information/education/cognitive ability

When people are not interested in the topic

When they are not motivated to answer carefully

When they get tired (late in the survey)

How do you solve this problem?


- The best way is to vary the order of responses randomly.
 - Half of all respondents get the responses with “plenty”.
 - Then statistically control for response order when you do analysis.

Thought Exercise

- Let's practice some of these ideas. Please take a few minutes and think about the **questions and response options on the handout**. Can you find anything that you should be changed about the response options in order to improve them? Write some notes in the margins, and then we can discuss your suggestions.


Any problems?


- 1. Based on what you know and have seen of Qatar Museums Authority's work, do you strongly agree, somewhat agree, somewhat disagree or strongly disagree with the statement that Qatar Museums Authority programs and activities represent Qatar properly?
-
- 1. Strongly agree
- 2. Somewhat agree
- 3. Somewhat disagree
- 4. Strongly disagree


- 
- 1. Many people may never have heard of the Qatar Museum Authority or their work. A lack of “Don’t Know” might be a problem.
 - 2. Agree-Disagree format will promote acquiescence bias.
 - 3. No “middle category”.

Problems?

- 2. Overall, how happy were you with your last visit to the Museum of Islamic Art?
-
- 1. Strongly approve
- 2. Somewhat approve
- 3. Somewhat disapprove
- 4. Strongly disapprove
- 5. Don't Know


- 
- 1. The response options (approve) do not match the dimension of evaluation (happy).
 - 2. This could increase “don’t know” responses, which is explicitly offered here. We probably do not need a “don’t know” option here if we fix problem 1.
 - 3. There is no middle option, and some people may neither approve nor disapprove.


- 
- 3. If you were given a choice would you like to see more emphasis on Arabic or on English at schools?
 - 1. English
 - 2. Arabic


- 
- There should be some middle categories here, since folks could want some degree of both.
 - 1. Strongly emphasize English
 - 2. Somewhat emphasize English
 - 3. Equal emphasis on both
 - 4. Somewhat emphasize Arabic
 - 5. Strongly emphasize Arabic

Problems?

- 4. How strongly could you support the marriage of a close family member to someone of a different nationality?
-
- 1. Very strongly
- 2. Somewhat strongly
- 3. Not too strongly
- 4. Not strongly at all
- 5. Don't Know

- 
- Here the “Don’t Know” response allows people to avoid seeming intolerant to people from different races or nationalities. We should leave it out.
 - It is also double barreled, because some might feel differently about race than about nationality.
 - Missing a category in the “middle” (such as “moderately”).

- 
- 5. Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?
 -
 - 1. Most people can be trusted
 - 2. Need to be very careful
 - 3. Don't Know

- 
- 1. This question has response options that are not mutually exclusive (you might think people can be trusted but you **still** need to be very careful).
 - 2. There is no middle category, and there are too few response options.
 - 3. Because of these problems, people might satisfice and choose “Don’t Know”.

Computer Exercise

Goal: Compare similar items with different response categories.

Run frequencies on the following pairs of variables:

1. traffic1 and traffic2
2. timenet1 and timenet2
3. know1_3 and know2_3

Compare how people answered each version of the question pairs.

Write a couple of sentences about which version of each question pair is best and why.

traffic1

Do you think that the current penalties for traffic violation in Doha are far too high, somewhat too high, somewhat too low, or far too low?

- Far too high
- Somewhat too high
- Somewhat too low
- Far too low

traffic2

Do you think that the current penalties for traffic violation in Doha are far too high, somewhat too high, about right, somewhat too low, or far too low?

- Far too high
- Somewhat too high
- About right
- Somewhat too low
- Far too low

Frequencies for traffic1 and traffic2

traffic1 Do you think that the current penalties for traffic violation in Doha are far too high, somewhat too...

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Far too high	4	14.8	23.5	23.5
	2 Somewhat too high	8	29.6	47.1	70.6
	4 Somewhat too low	3	11.1	17.6	88.2
	5 Far too low	2	7.4	11.8	100.0
	Total	17	63.0	100.0	
Missing	System	10	37.0		
Total		27	100.0		

traffic2 Do you think that the current penalties for traffic violation in Doha are far too high, somewhat too...

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Far too high	1	3.7	10.0	10.0
	2 Somewhat too high	5	18.5	50.0	60.0
	3 About right	3	11.1	30.0	90.0
	4 Somewhat too low	1	3.7	10.0	100.0
	Total	10	37.0	100.0	
Missing	System	17	63.0		
Total		27	100.0		

timenet1

Typically speaking, how long do you spend on the Internet each day?

- A few minutes - less than 1 hour
- 1 hour - less than 2 hours
- 2 hours - less than 3 hours
- 3 hours - less than 4 hours
- 4 hours or more

timenet2

Typically speaking, how long do you spend on the Internet each day?

- I do not use the internet
- A few minutes - less than 2 hours
- 2 hours - less than 4 hours
- 4 hours - less than 6 hours
- 6 hours or more

Means for timenet1 and timenet2

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
timenet1 Typically speaking, how long do you spend on the Internet each day?	12	2	5	4.00	1.206
timenet2 Typically speaking, how long do you spend on the Internet each day?	15	2	5	4.07	.884
Valid N (listwise)	0				

Frequencies for timenet1 and timenet2

timenet1 Typically speaking, how long do you spend on the Internet each day?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2 1 hour - less than 2 hours	2	7.4	16.7	16.7
	3 2 hours - less than 3 hours	2	7.4	16.7	33.3
	4 3 hours - less than 4 hours	2	7.4	16.7	50.0
	5 4 hours or more	6	22.2	50.0	100.0
	Total	12	44.4	100.0	
Missing	System	15	55.6		
Total		27	100.0		

timenet2 Typically speaking, how long do you spend on the Internet each day?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2 A few minutes - less than 2 hours	1	3.7	6.7	6.7
	3 2 hours - less than 4 hours	2	7.4	13.3	20.0
	4 4 hours - less than 6 hours	7	25.9	46.7	66.7
	5 6 hours or more	5	18.5	33.3	100.0
	Total	15	55.6	100.0	
Missing	System	12	44.4		
Total		27	100.0		

know1_3

Now some questions about Qatar. For these questions, we are going to show you some statements. For each one, could you please tell us whether you believe they are true or false? If you don't know, just say so.

	True	False	Don't know
There are thirty members of the Arab League.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Municipal elections are held every five years in Qatar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The Minister of Health is Abdulla Al-Qahtani.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

know2_3

Now some questions about Qatar. For these questions, we are going to read out some statements. For each one, could you please tell us whether you believe they are true or false?

	True	False
There are thirty members of the Arab League.	<input type="radio"/>	<input type="radio"/>
Municipal elections are held every five years in Qatar.	<input type="radio"/>	<input type="radio"/>
The Minister of Health is Abdulla Al-Qahtani.	<input type="radio"/>	<input type="radio"/>

Frequencies for know1_3 and know2_3

know1_3 Now some questions about Qatar. For these questions, we are going to show you some statements. For e...-The Minister of Health is Abdulla Al-Qahtani.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 True	7	25.9	53.8	53.8
	3 Don't know	6	22.2	46.2	100.0
	Total	13	48.1	100.0	
Missing	System	14	51.9		
Total		27	100.0		

know2_3 Now some questions about Qatar. For these questions, we are going to read out some statements. For e...-The Minister of Health is Abdulla Al-Qahtani.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 True	11	40.7	84.6	84.6
	2 False	2	7.4	15.4	100.0
	Total	13	48.1	100.0	
Missing	System	14	51.9		
Total		27	100.0		

Bibliography

Fowler, F.J. (2004) *Survey Research Methods, Fourth Edition*. Thousand Oaks, CA: Sage.

Fowler, F.J. (1995) *Improving Survey Questions*. Thousand Oaks, CA: Sage.

Groves, R. M., Fowler, F.J., Couper, M. P., Lepkowski, J.M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology, Second Edition*. New York: John Wiley.

Kalton, G. , M. Collins, and L. Brook. (1978). Experiment in Wording Opinion Questions. *Journal of the Royal Statistical Society, ser. C*, 27:149-61

Krosnick, Jon A. 1991. Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. *Applied Cognitive Psychology* 5: 213-36.

Krosnick, J. A., & Presser, S. (2010). Question and Questionnaire design. In J. D. Wright & P. V. Marsden (Eds.), *Handbook of Survey Research (Second Edition)*. West Yorkshire, England: Emerald Group.

Pasek, J., Schneider, D., Krosnick, J. A., Tahk, A., & Ophir, E. (Under Review). Prevalence and Moderators of the Candidate Name-Order Effect: Evidence from All Statewide General Elections in California.

Shuman, H., and S. Presser. (1981) *Questions and Answers in Attitude Surveys: Experiments in Question Form, Wording and Context*. New York: Academic Press.

Tourangeau, R., Rasinski, K., & Bradburn, N. (1991). Measuring Happiness in Surveys: A Test of Subtraction Hypothesis. *Public Opinion Quarterly*, 55-255-266.