Class 2: Principles of Questionnaire Design

MAST5953: Creating Your Own Data

Dr. Miriam Sorace

www.miriamsorace.net

17 November 2021

Quick Recap: Last Class

- Course structure and website
- Different types of surveys
  - Structured
  - Semi-structured
  - Unstructured
- How to write a report using RMarkdown
  - Downloading files from GitHub
  - Knitting the RMarkdown file
  - Do check out the RMarkdown Cheat Sheet and practice with it on your own!

Outline of Today's Class

Validity & Reliability

Formulating Questions

- Consider the following survey item examples, aimed at measuring social media use and detecting social media over-use:
  1. *"How many hours did you spend on Twitter yesterday?"*
  2. *"How many hours per day have you spent - on average - on social media sites in the past year?"*
- Which alternative would you pick? Would either one of them work? If not, why do you think the survey item is not appropriate to measure the level of social media use?

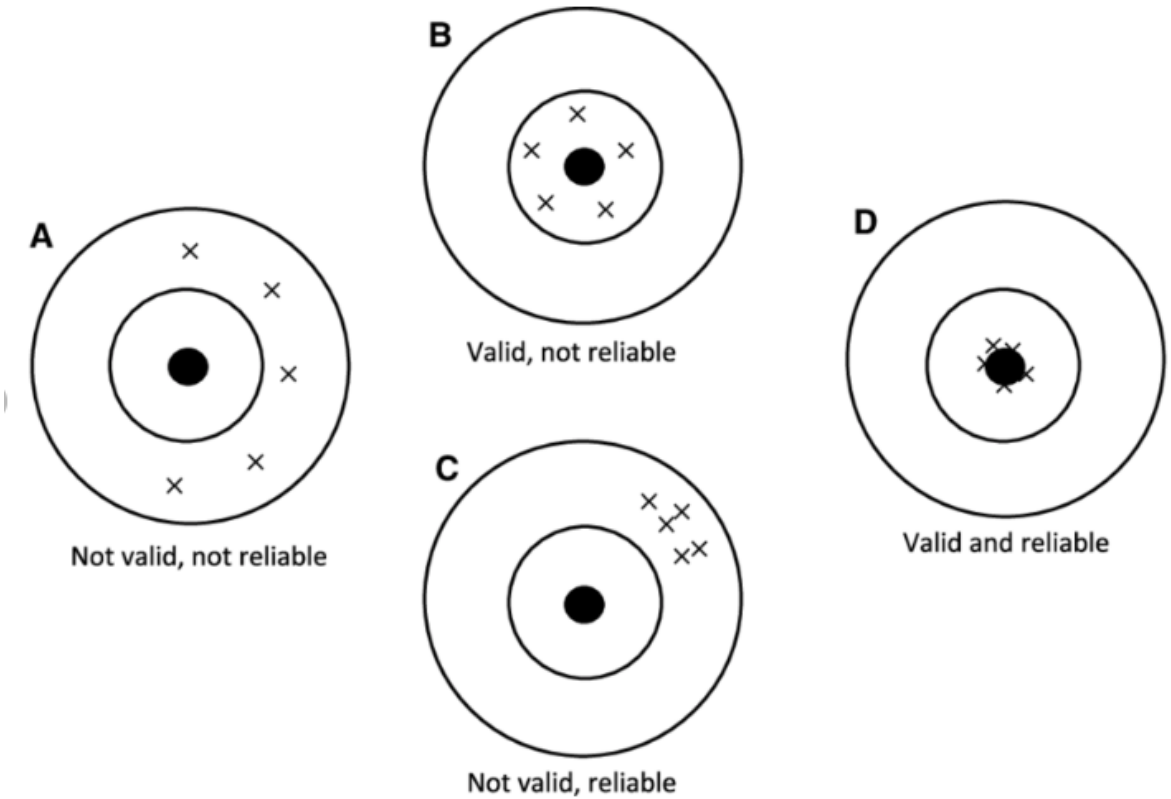
Validity

- The measure (i.e. the survey item) is related to the underlying construct.
  - E.g. the response delivers the true characteristic/behaviour/attitude that the researcher intended to elicit.

Reliability

- The measure (i.e. the survey item) performs consistently
  - E.g. the response to the question is consistent within the same individual over-time (*test-retest*), or different individuals respond using the same response process (*test-test*).

Validity vs. Reliability



This figure shows possible combinations of validity and reliability. Each 'target' contains a bullseye, which represents the real value of the measured property. A measure is reliable if repeated measures (×) fall within the same small area on the target. A measure is valid if repeated measures are clustered around the bulls-eye. A measure is more accurate the closer the x's are to the bulls-eye. a Shows an unreliable and also invalid measure. b Shows an unreliable but valid measure. c Shows a reliable but invalid measure. d Shows a measure that is both valid and reliable. Unreliable measures are either not valid (a) or inaccurate (b). However, reliable measures are not always valid (C)

1

[1] Image taken from: Irvine, Elizabeth. (2019). Developing Dark Pessimism Towards the Justificatory Role of Introspective Reports. *Erkenntnis*

Assessing Validity & Reliability

Going back to the examples

- Underlying construct = social media (over)use:

  1. *"How many hours did you spend on Twitter yesterday?"*

  2. *"How many hours per day have you spent - on average - on social media sites in the past year?"*

Testing Validity

- Correlation between the survey answers and some measure of the *actual* underlying construct
- Two main methods:
  - **Gold Standard**: data external to the survey
    - E.g. health responses vs. actual medical records
  - **Theoretical Correlate**: use responses to attitude/demographics that should be theoretically related
    - E.g. ideology question correlated with party ID self-report

Testing Validity

Applied Example: Political Knowledge

1. What is political knowledge?
2. How would you measure it with a survey?

Testing Validity

Applied Example: Political Knowledge Items [2]

1. Expert survey (random sample, 111 political scientists) on what should qualify as essential political knowledge to have.
   - Institutions/decision-making process (89%)
   - Policies/issues (82%)
   - History (77%)
   - Ideological alignments (62%)
2. Correlations with:
   - theoretical predictors: political efficacy, ideological stability, political participation & opinionation.
   - Correlation with interviewers' own rating of respondents' political knowledge (!)
   - Gold standard measures of pol. knowledge.

[2] Carpini, M. X. D., Keeter, S. (1993). Measuring political knowledge: Putting first things first. American Journal of Political Science, 1179-1206.

Testing Validity

Applied Example: Political Knowledge Items [3]

1. What are we testing? Is recall (declarative memory) knowledge?
2. Are we capturing lack of effort or lack of knowledge?
   - Is someone with better memory necessarily more politically knowledgeable/skilled?
3. Survey experiment
   - Control group: standard pol. knowledge battery
   - Treatment group 1: monetary incentive + standard battery
   - Treatment group 2: extra time + standard battery
4. 'Motivation to search' treatments increased correct answers!
   - Search skills are an important component of knowledge.
   - Existing measures underestimate political knowledge of respondents.

[3] Prior, M., Lupia, A. (2008). Money, time, and political knowledge: Distinguishing quick recall and political learning skills. American Journal of Political Science, 52(1), 169-183.

Testing Reliability

Main method:

- Compare results from repeated interviews (better for more stable constructs - e.g. personality)
  - E.g. same people interviewed twice or aggregate results compared from 2 different interviews

The Survey Response Process

The Survey Response Process

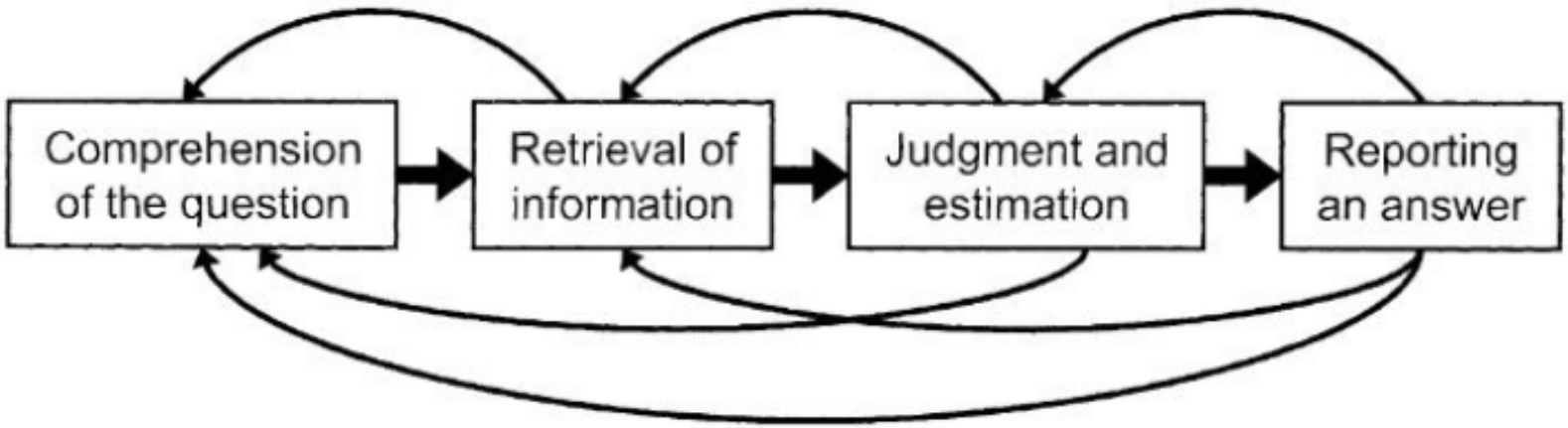


Figure 7.1 A simple model of the survey response process.

The Survey Response Process

- **Comprehension:** interpretation of the survey item
  - Parsing/reading
  - Assigning meaning to component concepts
  - Inferring the purpose 'behind' the question (pragmatics: inferences from co-text and context)
- **Retrieval:** recall of all pieces of information required to answer the question
- **Judgment:** combination/summary of the pieces of information recalled
- **Reporting:** formulation of the response in the required format

Key Terms/Concepts

- **Double-barreled question:** questions that link together two attitude objects). E.g. from Schuman & Presser (1981):
  - "Do you agree with the US government decision to defend Korea?"
  - vs. "Do you agree with the US government decision to defend Korea to stop a Communist take-over?"
- **Satisficing:** producing best answer available at lowest cognitive cost vs. 'optimizing' (answering comprehensively)
- **Social Desirability Bias:** The under-reporting of socially undesirable behaviours and over-reporting of socially desirable ones.
- **Positivity & Acquiescence Biases:** unwillingness to pick negative ends of the scale or disagree options, or extreme answer categories

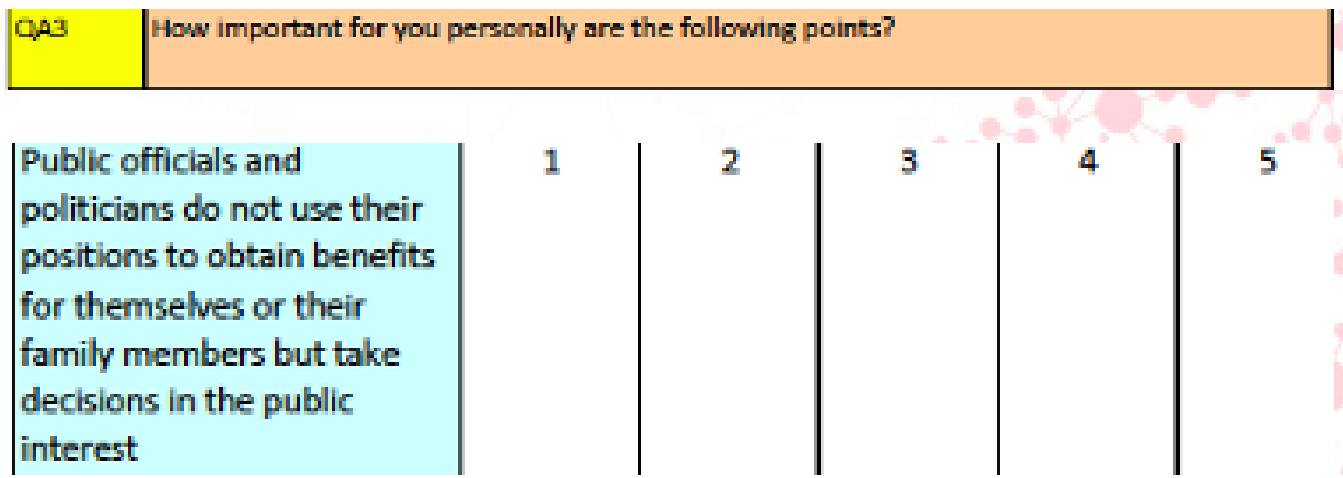
Evaluating Survey Items

Exercise: Survey Item Evaluation

4

[4] Taken from: https://yougov.co.uk/opi/surveys/results#/survey/1144947a-e4f8-11e9-ad60-47ee57021fb0

Exercise: Survey Item Evaluation



QA3	How important for you personally are the following points?				
Public officials and politicians do not use their positions to obtain benefits for themselves or their family members but take decisions in the public interest	1	2	3	4	5

5

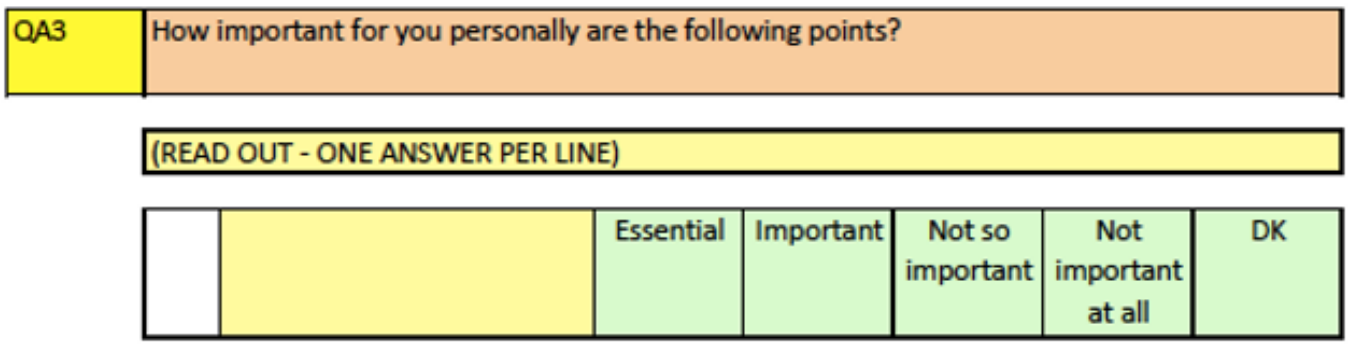
[5] Taken from: https://search.gesis.org/research_data/ZA7572?doi=10.4232/1.13372

Exercise: Survey Item Evaluation

How many children do you have?

- 1-2
- 3-4
- 5-6
- 6+

Exercise: Survey Item Evaluation



QA3	How important for you personally are the following points?					
	(READ OUT - ONE ANSWER PER LINE)					
		Essential	Important	Not so important	Not important at all	DK

[6]

[6] Taken from: https://search.gesis.org/research_data/ZA7572?doi=10.4232/1.13372

Solution: Pilot, Pilot, Pilot!

- "Survey researchers evaluate question comprehension, difficulty in memory retrieval, and related issues primarily by observing people trying to understand and answer the questions." (Groves et al. 2009: 259)

Solution: Pilot, Pilot, Pilot!

- Expert Reviews
- Focus Group Discussions
- Cognitive Interviews
- Field pre-tests
- Randomized 'split ballot' experiments
  - mostly used to pre-test different wordings of the same question

Cognitive Interviews

The standard pre-test

- Protocol analysis, introduced by Ericsson and Simon (1980)
- Interviewees 'think aloud' - or answer probing questions
- Alternatives to verbalisation:
  - Confidence ratings: respondents asked to assess confidence in their answer
  - Paraphrasing: respondents asked to re-formulate the question in their own words

Cognitive Interviews
Process

- Build useful **probes**
  - Questions/statements that encourage the respond to extend/elaborate on an answer. E.g. (1) Non-directive probing ("tell me more about it . . . "; "what do you mean by that?") OR (2) Summary technique (mostly in semi-structured and unstructured interviews)
- Those that follow Tourangeau's (1984) Survey Response Model are best:
  - "what does [concept x] mean to you?" (*Comprehension*)
  - "what information/knowledge did you collect/search in your memory to answer this question?" / "How did you decide which answer to give?" (*Retrieval*)
  - "how hard was this to answer?" / "was this question hard or easy to answer? Why?" (*Judgment*)
  - "Were the response categories adequate/sufficient?" (*Response*)

Cognitive Interviewing: Applied Example

From WDI UMich: Cognitive Interview Guide

Cognitive Interviewing
In-Class Practical Exercise

- Please click HERE and complete the survey
- I have placed the link in the GitHub page as well (under Class 2) - you can complete it on your phones/tablets as well!

Cognitive Interviews

Exercise - Cognitive Probing NSS Questions

(A) "I feel part of a community of staff and students"

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don't Know

1. What does 'community of staff and students' mean to you?
2. Can you take me through the steps of how you came to that answer?
3. Was the question easy or hard to answer?

(B) "Marking and assessment has been fair"

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don't Know

4. What does 'fair marking/assessment' mean to you?
5. Can you take me through the steps of how you came to that answer?
6. Was the question easy or hard to answer?

Cognitive Interviews

Exercise 2 - Your Own Cognitive Probing

- Think about your own survey question ...
  - Is there some phenomenon on which you'd like to measure public opinion/attitudes (support/behaviour frequency/importance)? Which is it?
- In pairs, present the survey question you thought of to your colleague and ask some cognitive probing questions.
- Just as a discussion at this stage, but take note of each other's comments.

Cognitive Interviews
Reporting Results

- % Don't Knows
- Average time taken to answer the question
- Comprehension probes: thematic analysis, e.g. hand-coding on the basis of a codebook
- Recall probes: count number of steps mentioned & calculate average.
- Difficulty probes: % saying "hard".

Cognitive Interviews

Best Practice. From: WDI UMich: Cognitive Interview Guide

- Inform the participant about confidentiality and anonymity procedures
- Respect participants' time: do not keep participants past the agreed time.
- Do not force participants to answer, and reassure that there is no right/wrong answer.
- Clarifications/question repetition might be requested. Note how often this happens, valuable data!
- Always thank the participant for their time!

Borrowing from Gold Standard Surveys

Don't re-invent the wheel!

- You can check survey items previously run by other surveys here: `https://www.icpsr.umich.edu/web/pages/ICPSR/ssvd/`
- Or you can play around with the Survey Quality Predictor `http://sqp.upf.edu/`

For Next Time ...

- Refine your survey item: what do you want to capture? How can you approach validity?
- Design your own cognitive interview: what probing questions would you field?

What did we learn today?

- Validity & Reliability
- The Survey Response Process
  - Comprehension
  - Retrieval
  - Judgment
  - Reporting
- Biases in answering survey questions
- How to test for validity via cognitive interviews