Utrecht University

Summer Course Survey Research: Advanced Survey Design

Bella Struminskaya & Peter Lugtig

Department of Methodology & Statistics, Utrecht University

© Lugtig, Struminskaya, Utrecht University
Slides by De Leeuw, Schouten, Hox, Struminskaya, Lugtig

Utrecht University

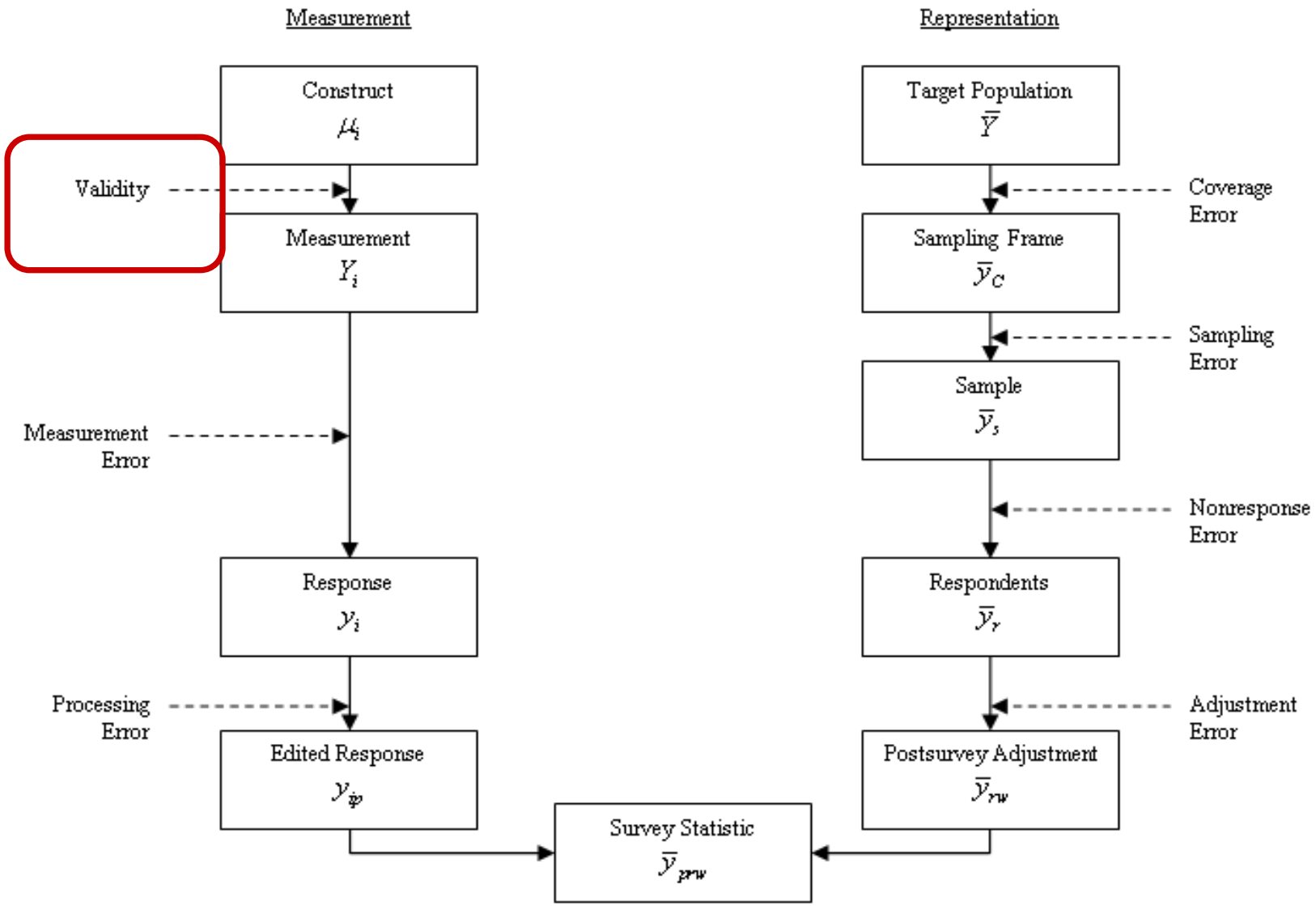
Advanced questionnaire design

Bella Struminskaya & Peter Lugtig

Department of Methodology & Statistics, Utrecht University

Operationalization

- Translating your research question into (prototype) questions
- Need problem analysis
  - Background of research
  - Research objective
  - Conceptual model
  - Intended use of research outcomes
- Has all to do with
  - Reducing specification (= measurement) error

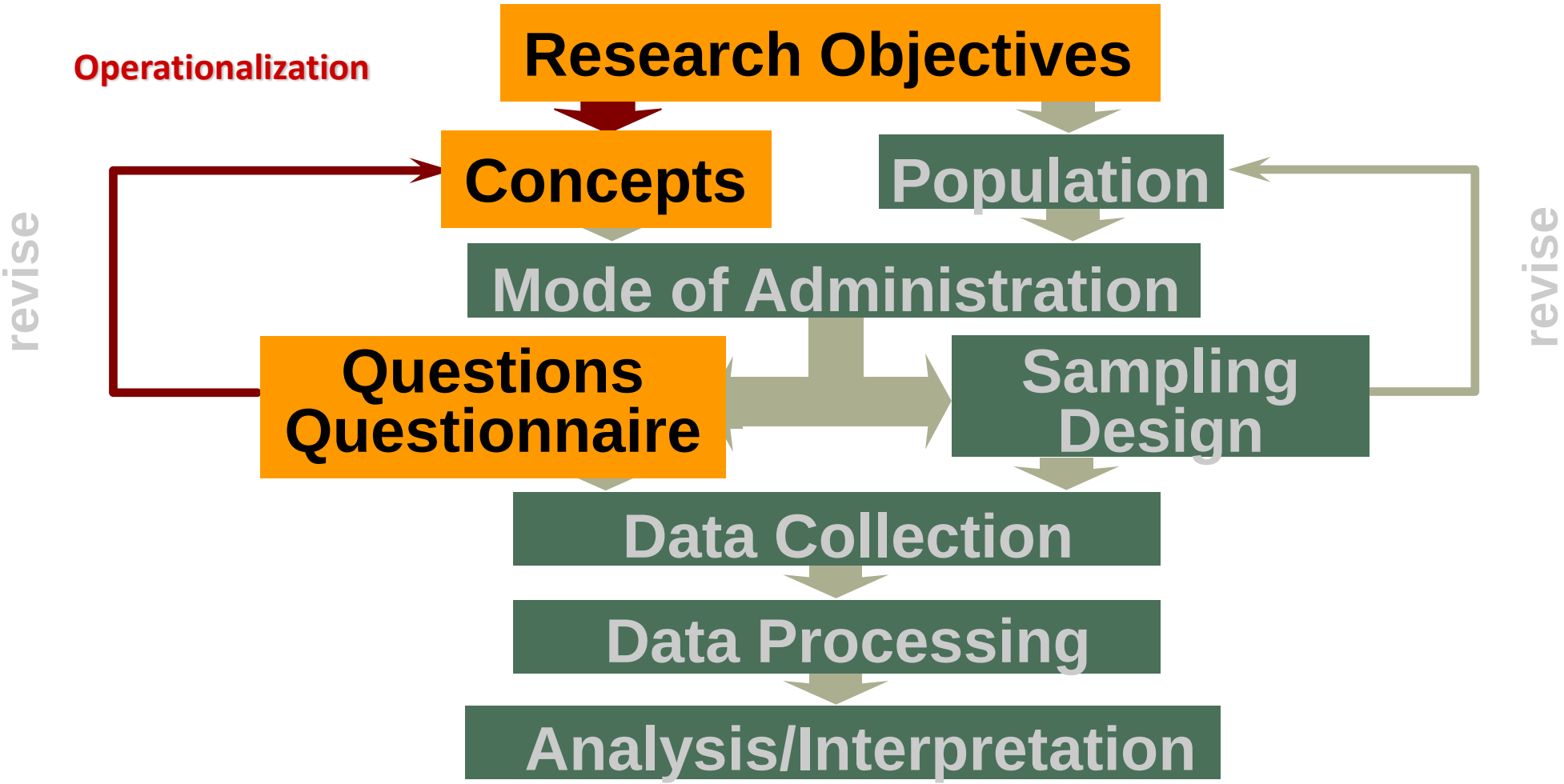


(Groves et al. 2009, p.48)

Specification Error

- Specification errors occur when the ***survey*** question fails to ask respondents about what is essential to answer the ***research*** question
- The concept implied in the question differs from the concept that should be measured (intended concept)
  - ***Validity problem***
- Wrong parameter is estimated!
  - The research question is not answered

The Survey Process



Questionnaire Step-by-Step

- List the research questions
- For each research question list
  - Important concepts; survey question topics
- List all required auxiliary information
  - Background variables needed for planned analyses, for planned comparisons
- Make first attempt: Write draft questions
- Consider the answer categories
- Draft introduction and instructions
- Revise questions, check phrasing
- Propose question order
- Pretest, informal methods
- Revise, optimize for data collection method, test

Construct & Question 1

- “For which party do you intend to vote?”
  - The question is simple
  - Our interest is the implications for the next election
  - Maybe we should ask a more complicated question, such as how sure respondents are of their answer, or which other parties they are considering

Construct & Question 2

- “What is your present occupation?”
  - Our goal is not to find out what people do at work
  - Our goal is to assign a score on Social Economic Status (SES)
- The survey question is simple, the theoretical concept SES is not
  - Different conceptualizations of SES may lead to different survey questions

From Concept to Survey Question

- Conceptualization
  - Make explicit what is intended, what one wants to measure
    - Delinquency? Deviance?
- Operationalization next step
  - From theoretical concept to proto-question
- Some concepts are very close to the actual survey questions: e.g., *age*
- Others are not, e.g., *poverty, SES, well-being, happiness*

From Concept to Question

- Starting with theoretical concept (Top Down)
  - Dimension/Indicator analysis
- Starting with Data (Bottom-up)
  - Qualitative in-depth interviews on topic

    → Relevant words, aspects
  - Inventory of existing questions / scales

    → Quantitative analysis

Top down

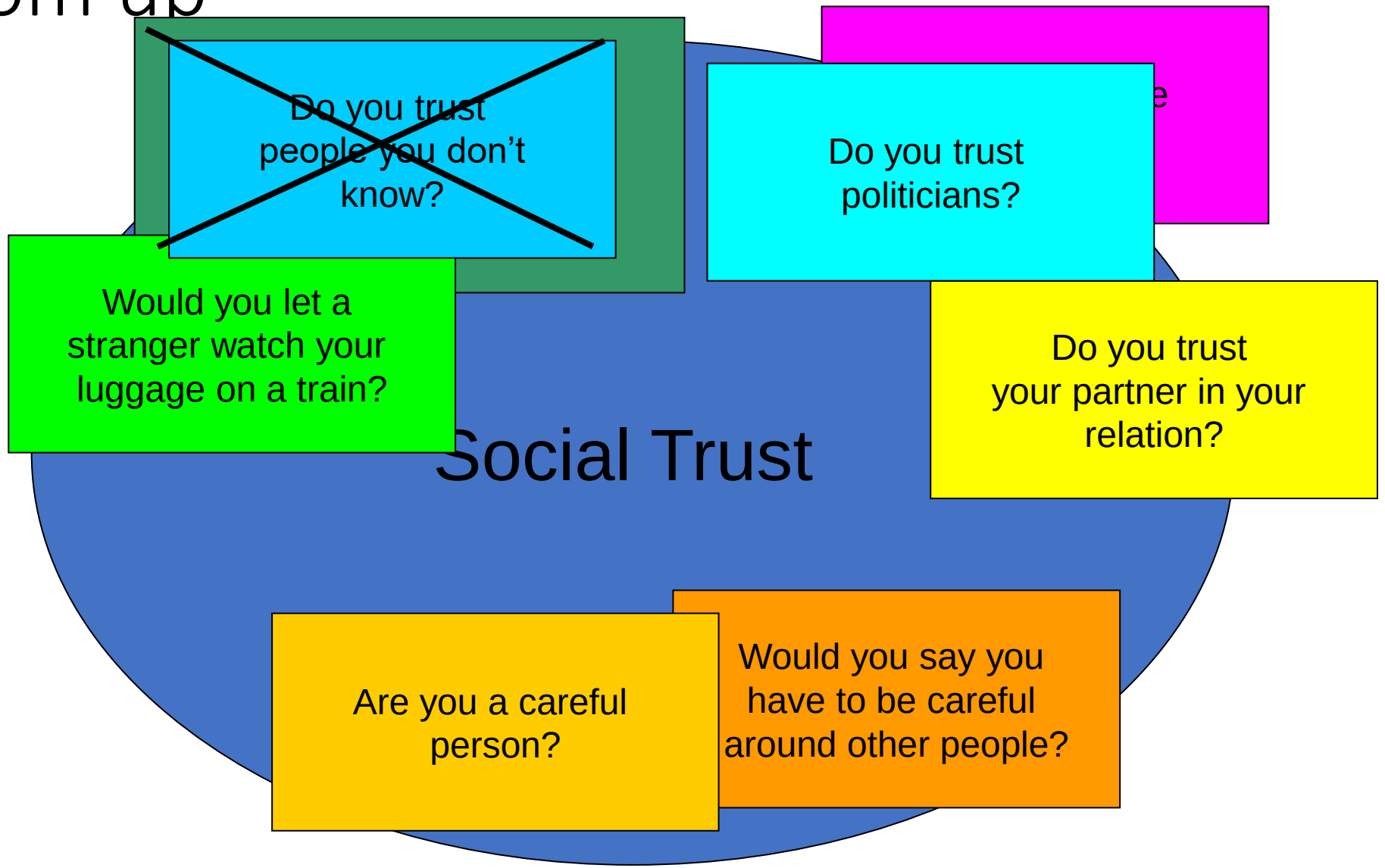
Say we want to measure ‘social trust’

In general: careful/reckless?
Trust in politicians?
Trust in strangers
Etc.

Operationalisation step by step

1. Use literature!!!
2. Identify different dimensions of constructs e.g. situational, persons or institutes
3. Come up with 1 (or a few) indicators per dimension
4. Write a survey question for every indicator
5. (think about how to combine all indicators into a total score)

Bottom up


Do you trust people you don't know?
Do you trust politicians?
Would you let a stranger watch your luggage on a train?
Do you trust your partner in your relation?
Social Trust
Are you a careful person?
Would you say you have to be careful around other people?

From Concept to Question

- Approaches compliment each other
- In good practice often mixed approach

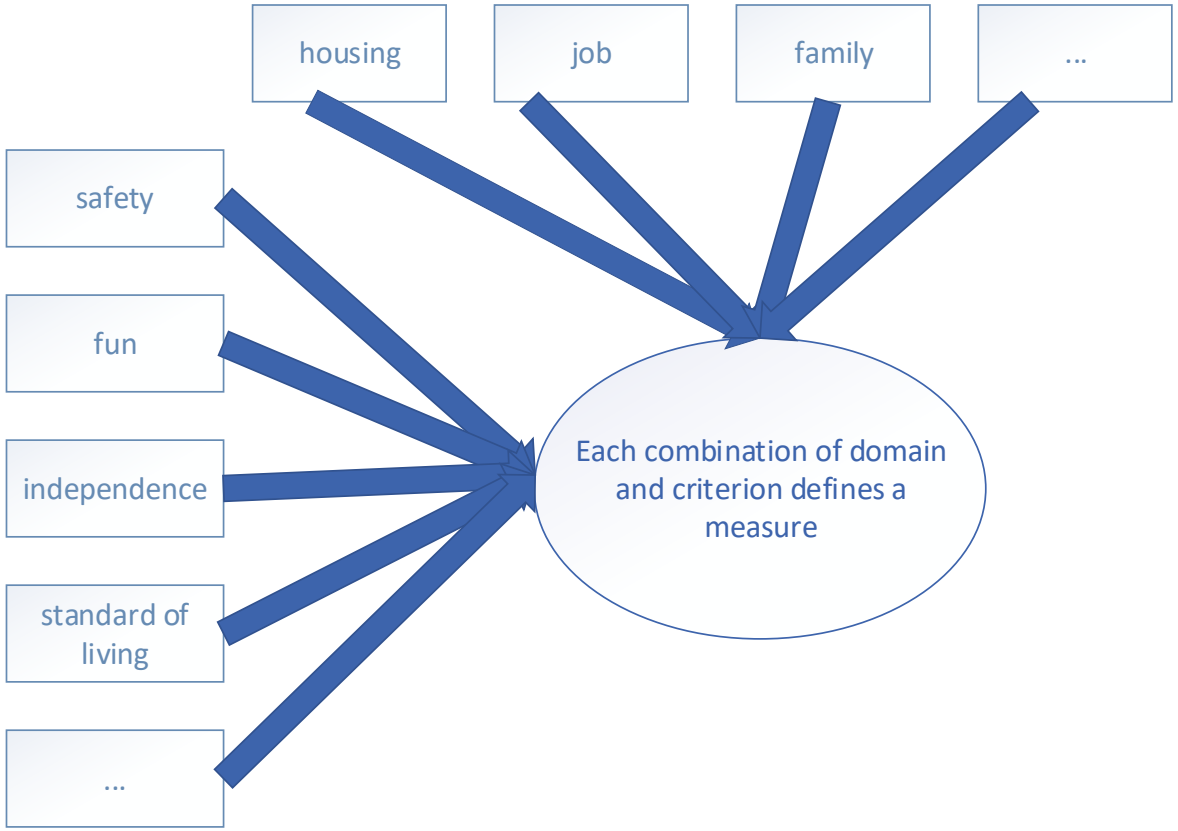
    e.g., Loneliness (Jenny Gierveld)

Top Down: Dimension/Indicator Analysis

- Steps
  1. Start with a vague concept
  2. Concept specification: Distinguish different components or dimensions
  3. Decide on type of information
  4. Specify empirical indicators for each dimension
  5. Write prototype questions for each indicator
- Important Checks
  1. Are all relevant phenomena included?
  2. Are the dimensions precise enough?
     - Too broad: divide in subdimensions
  3. Is it clear what is NOT included?
     - How does it differ from the other theoretical constructs?

Example: Well-being

- Long tradition of research into happiness & wellbeing (Andrews & Whitney)
  - Domains: housing, job, neighborhood, family, friends
  - Criteria: standard of living, safety, satisfaction, fun, etc.



Hox 1997

Examples Proto-Questions

- How would you feel about your house or apartment if you considered only the standard of living it enables you to have?
- How would you feel about your house and apartment if you considered only the fun it enables you to have?
- How would you feel about your house and apartment if you considered only it’s safety?
- Etc., etc.

Data Driven

- Define a vague and broad construct
- Collect many questions that seem relevant
  - Scour the literature for existing measures
- Judge which are relevant
- Do a special study:
  - Collect lots of data, investigate which questions appear to measure the same construct

    → Factor analysis / Scaling techniques
  - Assume that each factor (scale) measures one underlying theoretical construct
  - Cross-validate / Replicate

Example: Content Sampling

- Content analysis of 140 scales used in American sociology for well-being (Schuesler)
  - Most scales used rarely, few ‘classics’
  - Total 9500 questions, about 500 distinct, about 237 judged non-duplicates and potentially useful
  - These 237 questions put to sample of 1522 persons
  - Factor analysis leads to 12 well-being scales containing 95 items
- Other examples:
  - Cattel’s mood scales
  - Big Five

Practice: often a mix

- Content sampling adaptation
- Search literature
  → Collect questionnaires
- Judge questions
  → How relevant are they for your research?
- Use dimension analysis (or method of list of important indicators, or results of previous qualitative studies) to help you choose and decide
- Add questions you think are missing
  → For instance, through focus groups
- Test
- Get data and see if structure is there

Example: Loneliness

- Started Dictionary
- Reading stories (novels)
  - What is common, which terms are used
- Open Interviews and Focus groups
  - Examples, words, descriptions
- Content sampling
  - Other questionnaires
- Add what is missing, leave out what is not effective
  - Judging helped by dimension/indicator research

(J. Gierveld)

Questionnaire Design In Sum

- Operationalization:
  - What concepts do we like to measure
  - What are important aspects
    - Find empirical indicators for each aspect
    - Empirical indicators “not-quite-variables”
    - Empirical indicators are ‘proto-questions’
      (Not fully written questions with all response alternatives)
- Questionnaire development:
  - Write full questions
  - Order them
  - Write instructions/introductions
  - Test questions/instrument

Stages Questionnaire Design

- Specification topic
- Outline questionnaire
- Preliminary version questionnaire
- Pretest
- Revision

Practical Limitations

- Well-specified Concepts and Objective Survey

but

- Limited Length of Questionnaire
  - Select questions for questionnaire carefully!
- Needed
  - Knowledge research objectives
  - Planned Analyses

Some Words on Question Utility

1. Does question measure aspect of research question? *or*
2. Does question provide information needed in conjunction with other variable (analysis)?
3. Will most respondents understand question?
4. Will most respondents have information?
5. Will most respondents be willing to answer?
6. Should this question be asked to all or to a subset? *If so,*
7. Is subset identifiable?
8. Is other information needed to analyze the question? *If so,*
9. Add necessary questions (go-back to 2)

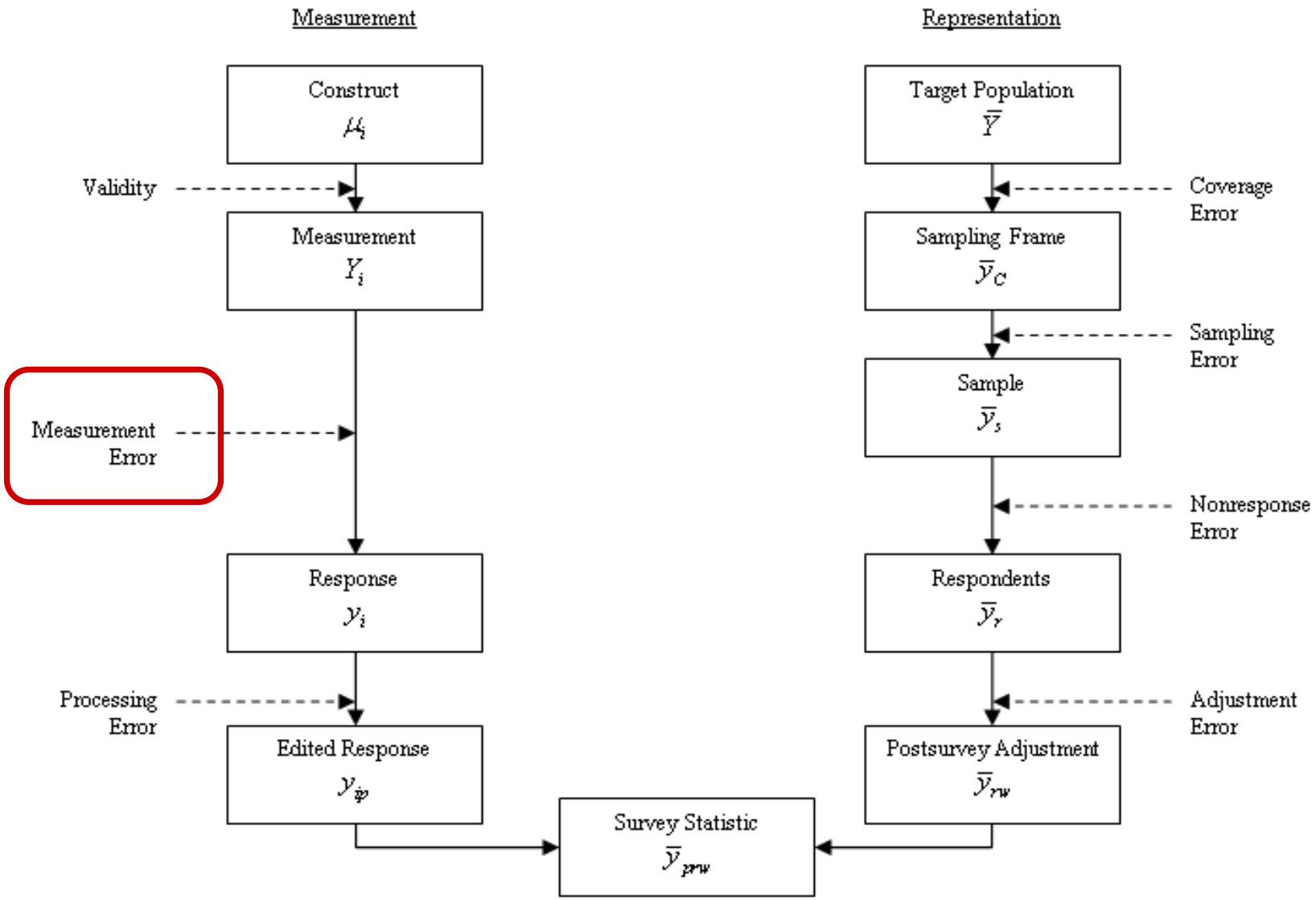
From: Czaja & Blair, 1996, 2005

Based On

- Chapter 8: Writing Effective Questions
  - E.D. De Leeuw, J.J. Hox & D.A. Dillman (2008) International Handbook of Survey Methodology. New York: Lawrence Erlbaum Associates, Taylor and Francis Group.
- Chapter 2: From theoretical concept to survey question by J. Hox
  - L. Lyberg, P. Biemer, M. Colins, E. de Leeuw, C. Dippo, N. Schwarz, D. Trewin (1997) Survey Measurement and Process Quality. New York: Wiley.

Further readings:

- Blair, J., Czaja, R. & Blair, E. (2014). Designing surveys: A guide to decisions and procedures, 3rd edition, Sage
- Fowler, F. J. (1995). Improving survey questions: Sage



(Groves et al. 2009, p.48)

How many hours per day do you typically watch TV?

Group A	Group B	Group C
½ hour or less	2½ hour or less	open-ended question
½ - 1 hour	2½ - 3 hours	
1 - 1½ hours	3 - 3½ hours	
1 ½ - 2 hours	3½ - 4 hours	
2 - 2 ½ hours	4 - 4 ½ hours	
more than 2 ½ hours	more than 4 ½ hours	

How many hours per day do you typically watch TV?

	Group A	Group B	Group C
	½ hour or less	2½ hour or less	open-ended question
	½ - 1 hour	2½ - 3 hours	
	1 - 1½ hours	3 - 3½ hours	
	1 ½ - 2 hours	3½ - 4 hours	
	2 - 2 ½ hours	4 - 4 ½ hours	
	more than 2 ½ hours	more than 4 ½ hours	
More than 2.5 hours	**22%**	**54%**	**52%**

Measurement error

- Deviaton of the answers of respondents from their true values
- Y=T+e
- Answer obtained=true score+error
- ME respondents: lack of motivation,comprehenson problems
- ME survey instrument: poor wording, poor design, technical flaws
- Random error → reliability
- Systematic error → validity
- Design issues: questionnaire design

Introduction text: questions that arise

- What is the study about?
- Who is conducting it?
- Why is the study important?
- Why (and how) was I (the respondent) selected?
- What will be done with the study results?
- Are my answers dealt with confidentially?
- How long will it take me to complete the survey?
- What’s in it for me?

Order of questions

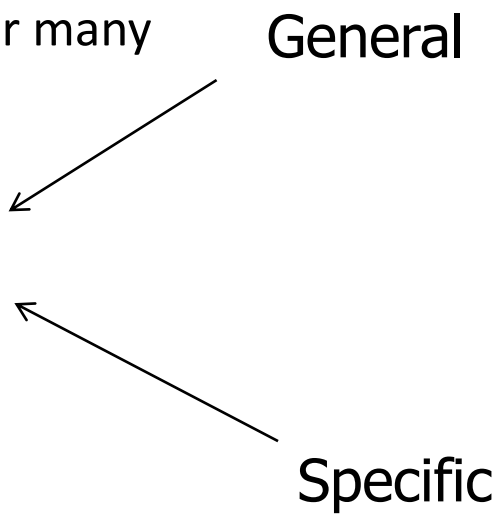
- Start with *salient and interesting* questions
- First questions should be related to the topic of the study
- A questionnaire is a constant conversation between the researcher and the respondent
  - It typically evolves in accordance with societal norms on conversations.
  - Each question should logically follow from the previous question (sometimes helped by question introductory texts).
  - Jumping quickly between topics means that answers are less likely to be well thought out, as new topics evoke top-of-the-head responses

Ordering affects survey results: from general to specific

- Answer to one question depends on the previous one(s)
- Common in item grids/matrices, when respondents have to answer many questions

General question **A**: How satisfied are you with your life in general?

Specific question **B**: How satisfied are you with your relationship?

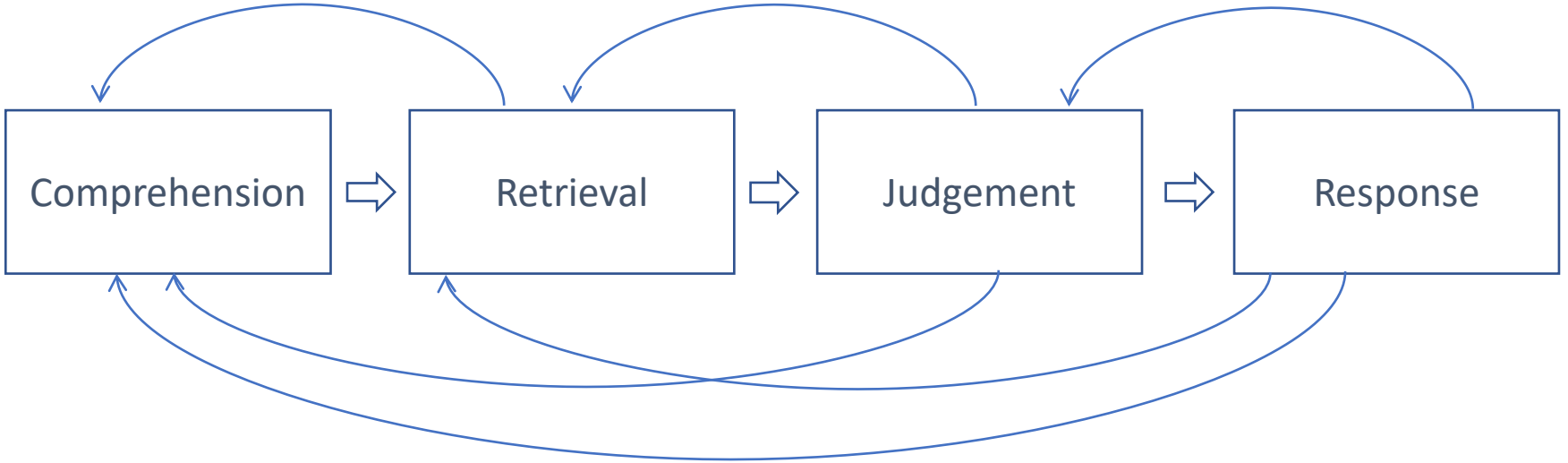


Correlation between questions differs, when ordering differs.

AB: r = 0.32 (questions interpreted as indicating to separate domains)

BA: r = 0.67 (questions interpreted as indicating connected domains)

Psychology of asking questions



Errors are possible at every stage of the process!

Not all steps may be taken (“satisficing”, see the work of Jon Krosnick)

Tourangeau et al. 2000: 8

Psychology of asking questions

Careful answers are the product of:

Step 1: Comprehension of the question

Respondent error: Misinterpret the question

Step 2: Retrieval of relevant information (cognitive processing)

Respondent error: Forget crucial information

Step 3: Use of information to make evaluation (generating an opinion)

Respondent error: Make erroneous inferences based on what is retrieved

Step 4: Selection and reporting of an answer (formatting the response)

Respondent error: Map answer onto an inappropriate response category

Typical response effects

Acquiescence: tendency to agree

Solution: Do you agree or disagree?

Solution: positive and negative worded items

Social desirability: need to present oneself in a favourable light

Solution: self-administered survey

Solution: randomized response technique

Primacy effects: more easily selecting the first answer (Note: primacy in web surveys compared to e.g. recency in aural modes)

Solution: randomization of response categories

Satisficing: searching for the first option that seems appropriate, rather than evaluating each option individually before arriving at an answer

Standards for survey questions

- **Content standards**
  - are the questions asking about the right things?
- **Cognitive standards**
  - do respondents understand the questions consistently;
  - do they have the information required to answer them;
  - are they willing and able to formulate answers to the questions?)
- **Usability standards**
  - can respondents complete the questionnaire easily and as they were intended to?

Groves et al. (2009)

Does the following question meets these standards:

[Please assume that you are not retired yet]

Next you will find four options for how you could spend your money over your lifetime. For each option the first column indicates how much you/ your household could spend on average per month from age 25 until retirement.

Thus, this refers to your total (working) time from age 25 until retirement.The second column indicates how much you / your household could spend during retirement.

Please think of all your expenditures, such as food, clothing, housing, insurance, traveling etc. Assume that the numbers below show what you can spend after having already paid for taxes. Assume also that prices of the things you spend your money on remain the same in the future as today (i.e., no inflation). If you had a choice, which option would you like most?

Principles for writing survey Questions 1: Dillman's Tailored Design Method (2007)

- Choose simple over specialized words
- Choose as few words as possible to pose the question
- Use complete sentences to ask questions
- Avoid vague quantifiers when more precise estimates can be obtained
- Avoid specificity that exceeds the respondent's potential for having an accurate, ready-made answer
- Use equal numbers of positive and negative categories for scalar questions

Principles for writing survey Questions 2: Dillman's Tailored Design Method (2007)

- Distinguish undecided from neutral by placement at the end of the scale
- State both sides of attitude scales in the question stems
- Eliminate check-all-that-apply question formats to reduce primacy effects
- Develop response categories that are mutually exclusive and exhaustive
- Use cognitive design techniques to improve recall

Principles for writing survey Questions 3: Dillman's Tailored Design Method (2007)

- Provide appropriate time referents
- Be sure each question is technically accurate
- Choose question wordings that allow essential comparisons to be made with previously collected data
- Avoid asking respondents to say yes in order to mean no
- Avoid double-barreled questions
- Soften the impact of potentially objectionable questions
- Avoid asking respondents to make unnecessary calculations

Answer categories

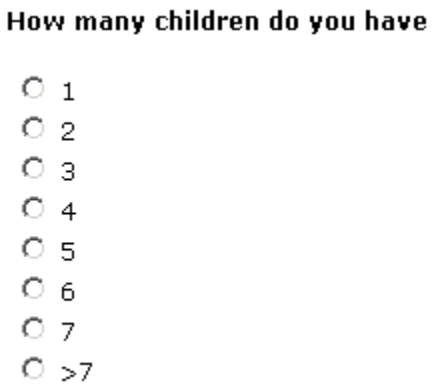
- Open-ended
  - Pro: no response category effect
  - Con: more effort for respondent and researcher (no guidance/rounding and recoding)
  - Open-ended items are preferable when the required form of the answer is
    - quite obvious (e.g. What is your age?)
    - when you want your respondent to elaborate on a topic (Why do you think so?)
    - or when it is difficult to provide a good set of response alternatives (What is your occupation?)
- Closed-ended
  - Ordered
  - Unordered

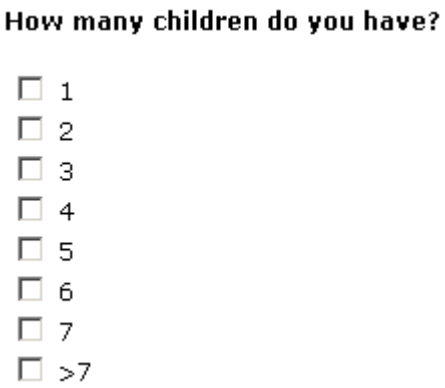
Ordered answer categories

- Scalar questions:
  - Strongly agree to strongly disagree
  - Very favorable to very unfavorable
  - Excellent to poor
  - Extremely satisfied to extremely dissatisfied.
- Issues with scalar questions:
  - a 5, 7, 9, or 11-point scale?
  - verbal labels or endpoint labels only?
  - Add numbers?
  - Add color?

Answer categories (in online surveys)

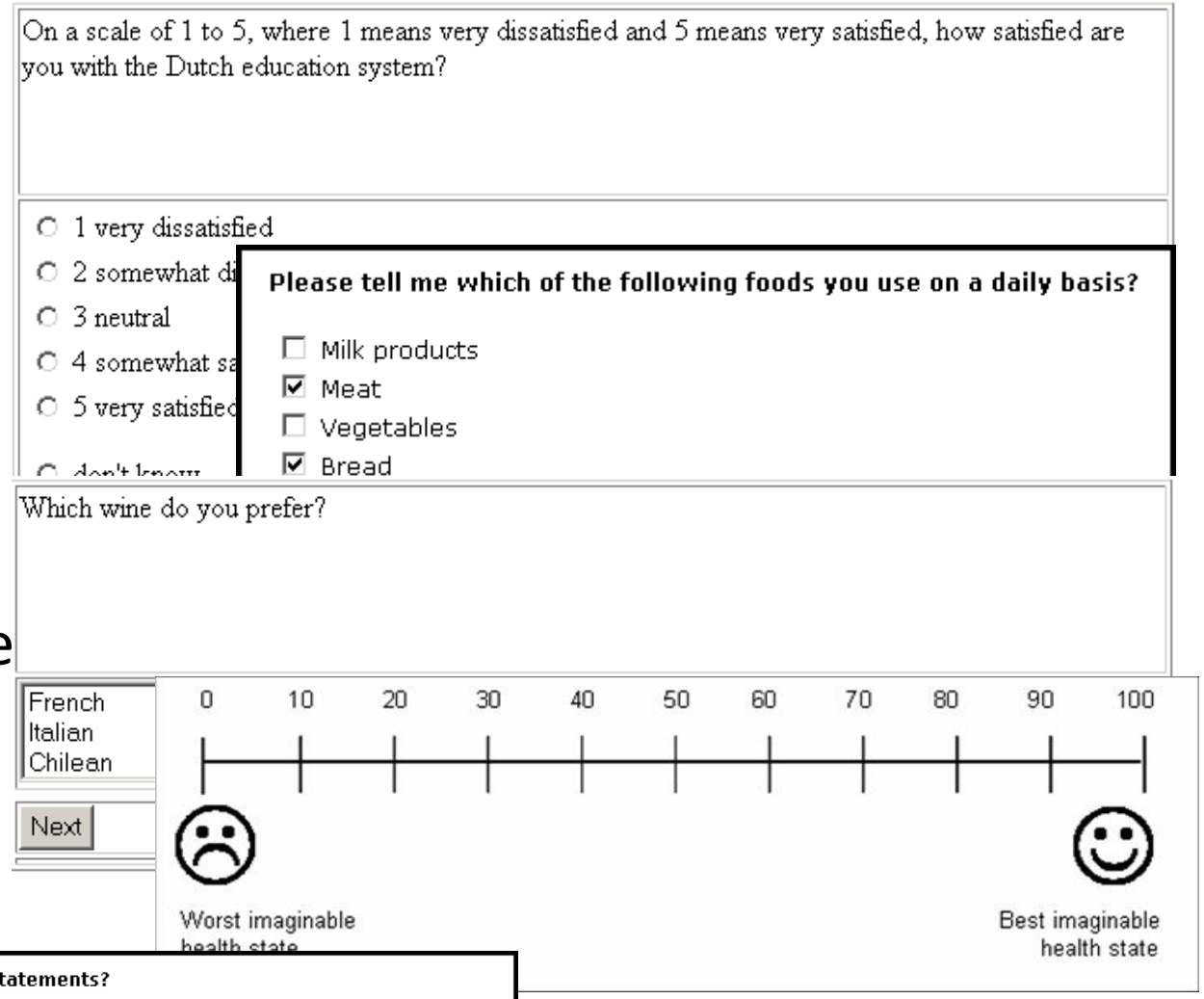
- Example question 1:
- How many children do you have?

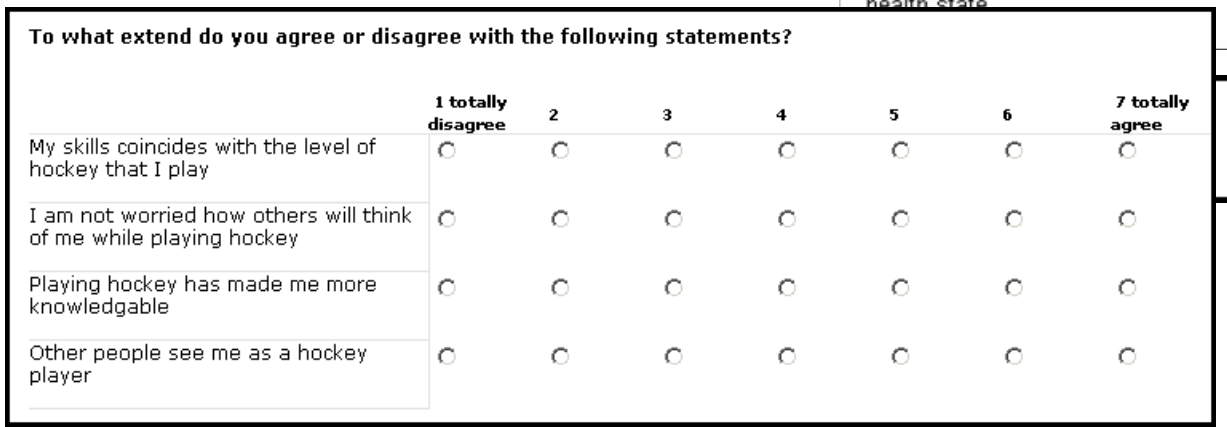


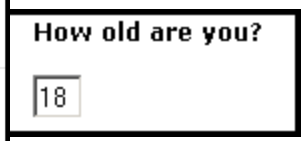


Answer formats

- Radio buttons
- Check boxes
- Dropdown menu
- Slider bars (visual analogue
- Text box
- Text area
- Matrix/Grid

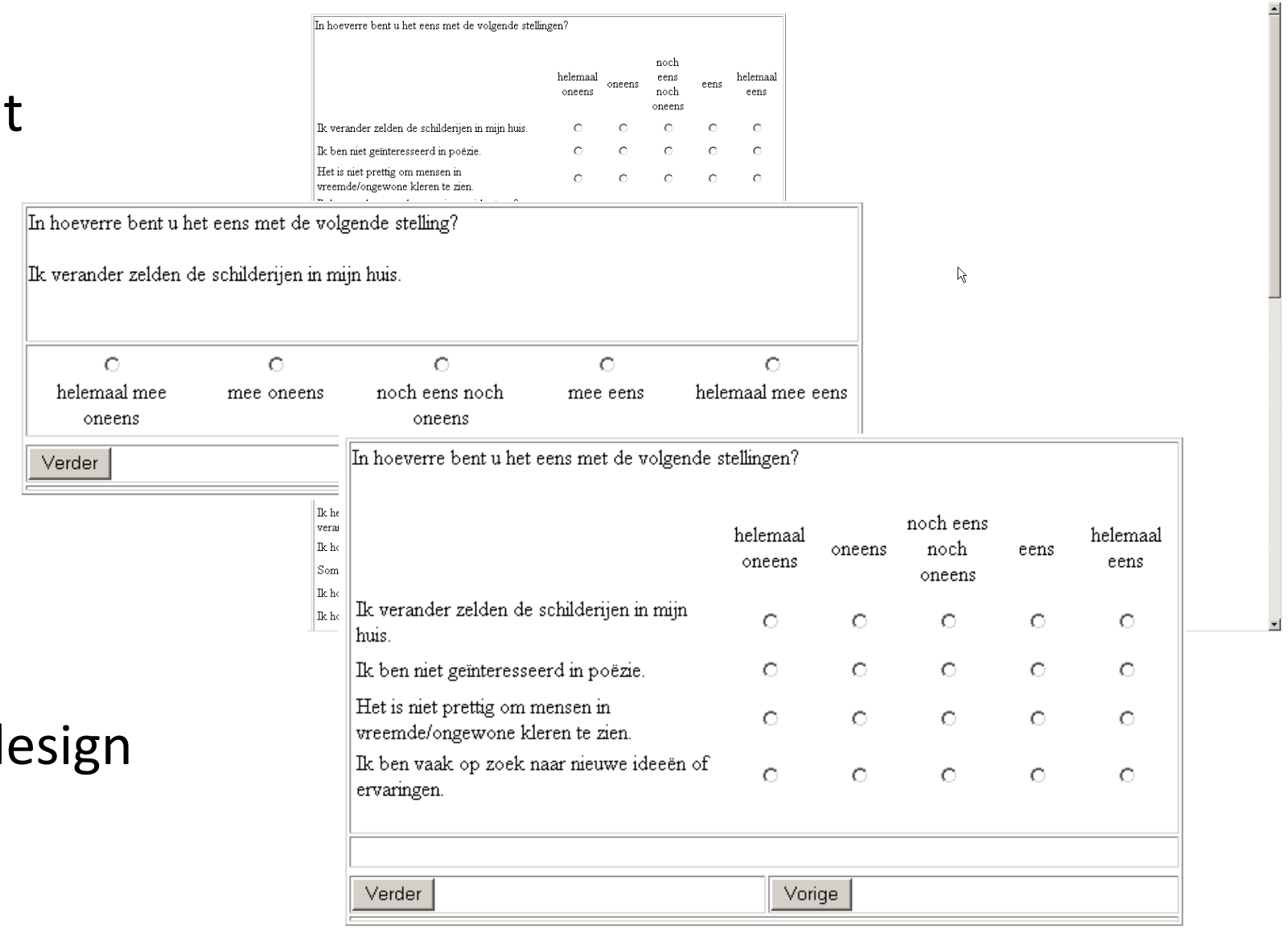






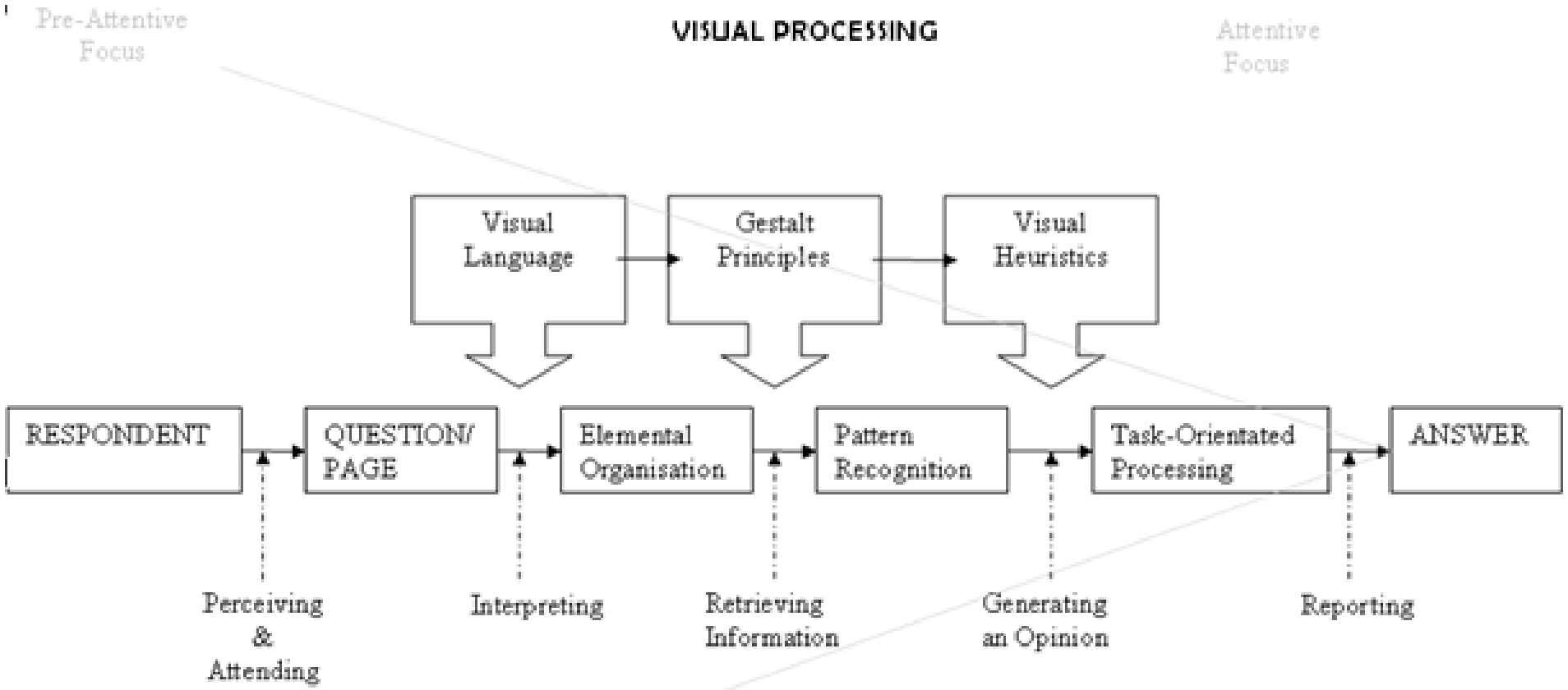
Programming the survey

- Scrollable format



- Page-per-page design

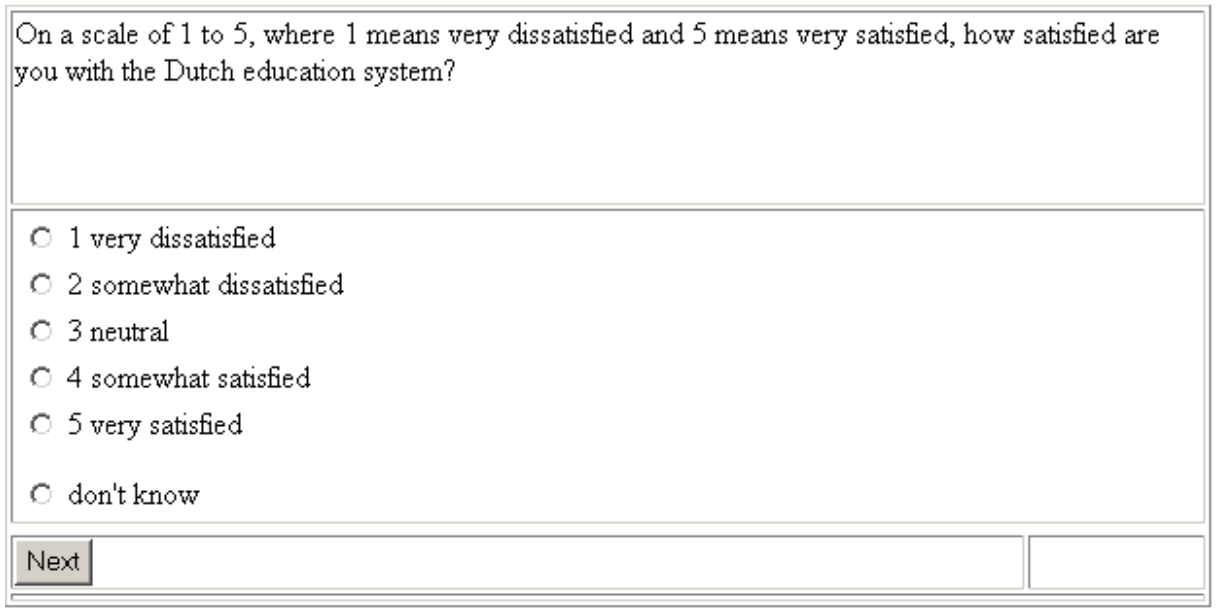
Visual Design



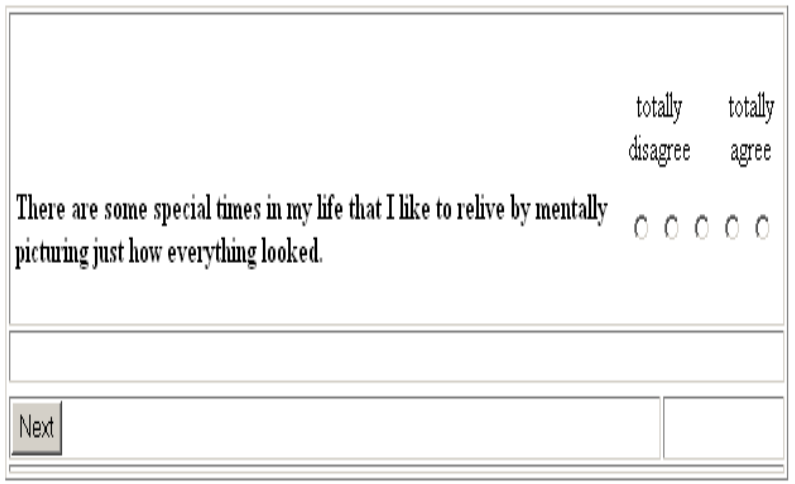
Toepoel &Dillman 2010

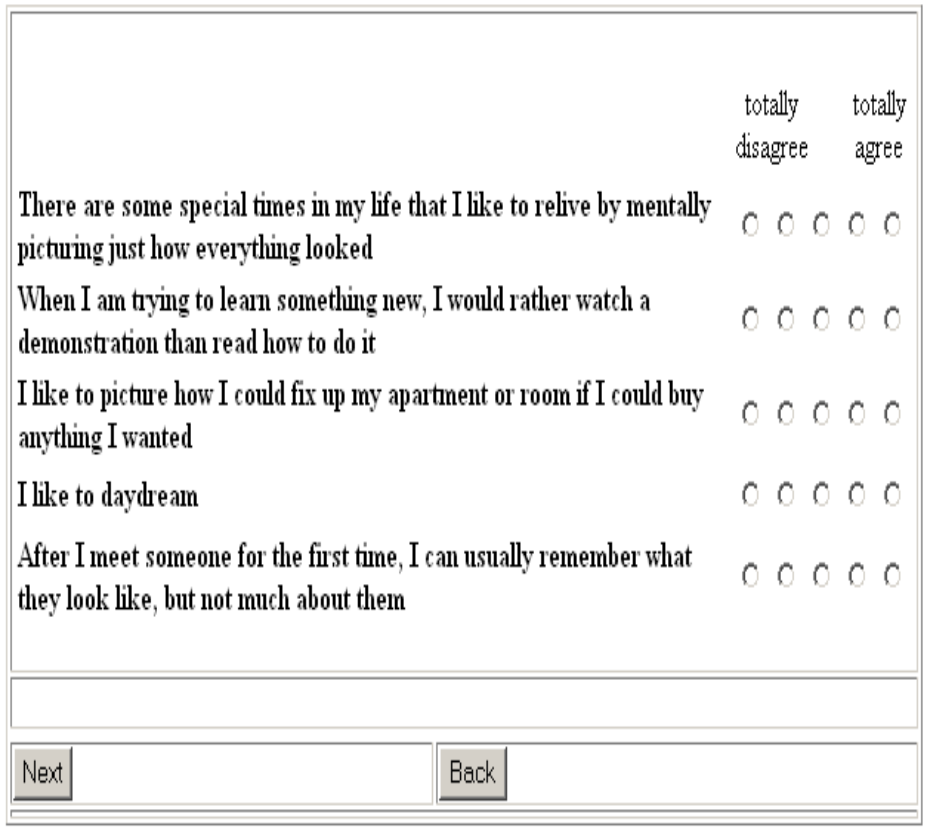
Visual heuristics

- Ignoring the “middle means typical” heuristic affects answers



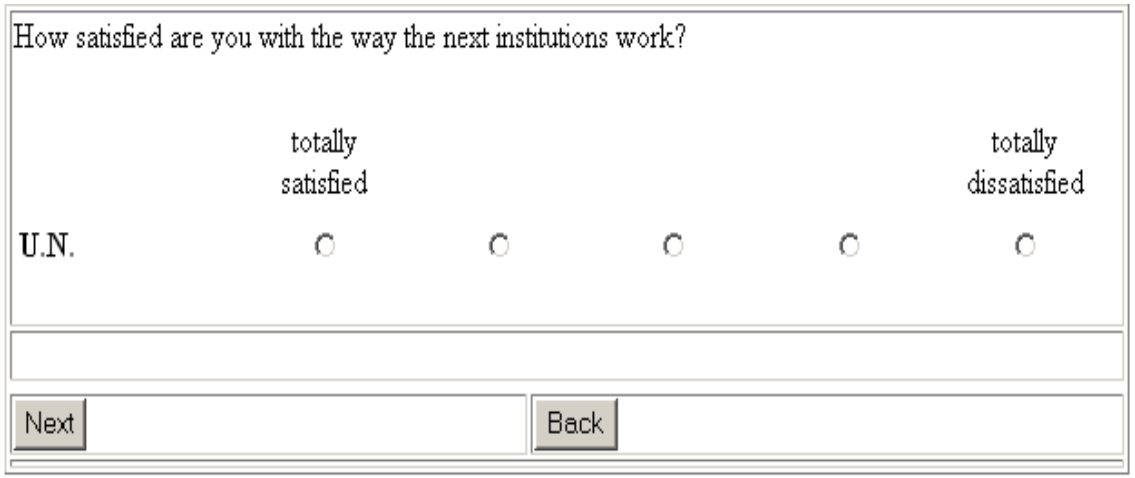
Ignoring “near means related” may affect answers

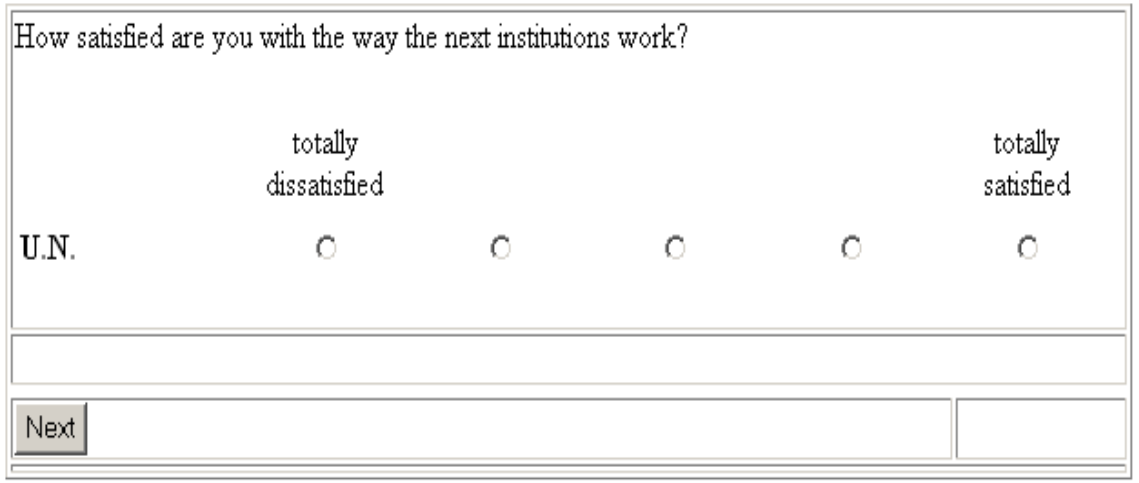




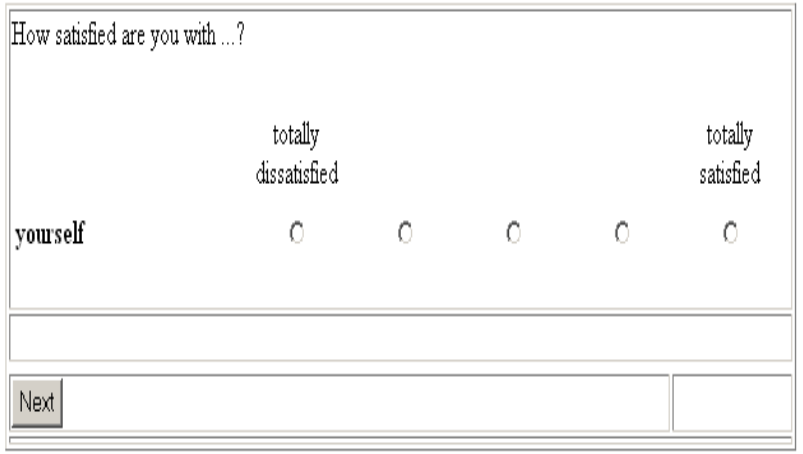
Ignoring “up/left means good” may affect answers

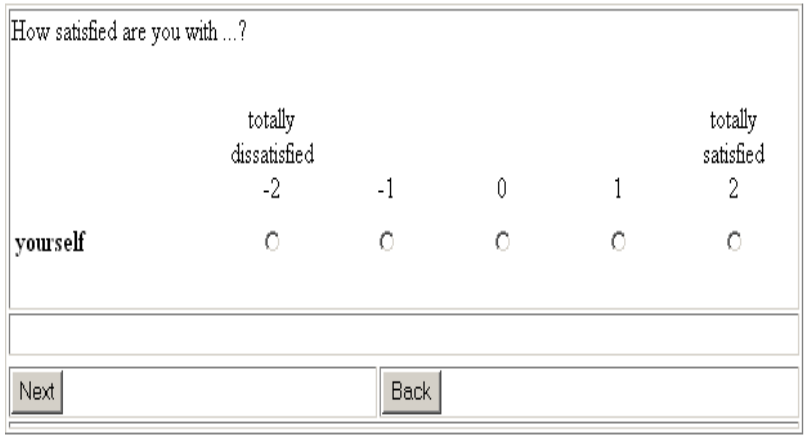
- Cultural differences English versus Dutch speaking countries





Ignoring “like means close” heuristic may affect answers





Unnecessary pictures influence answers

- “How often did you go out to eat in the last month?”

Also apparent in follow-up questions

- “How often did you go shopping in the last month?”
- Experiment: Respondents were more likely to go alone and enjoyed their trip less when confronted with the grocery picture

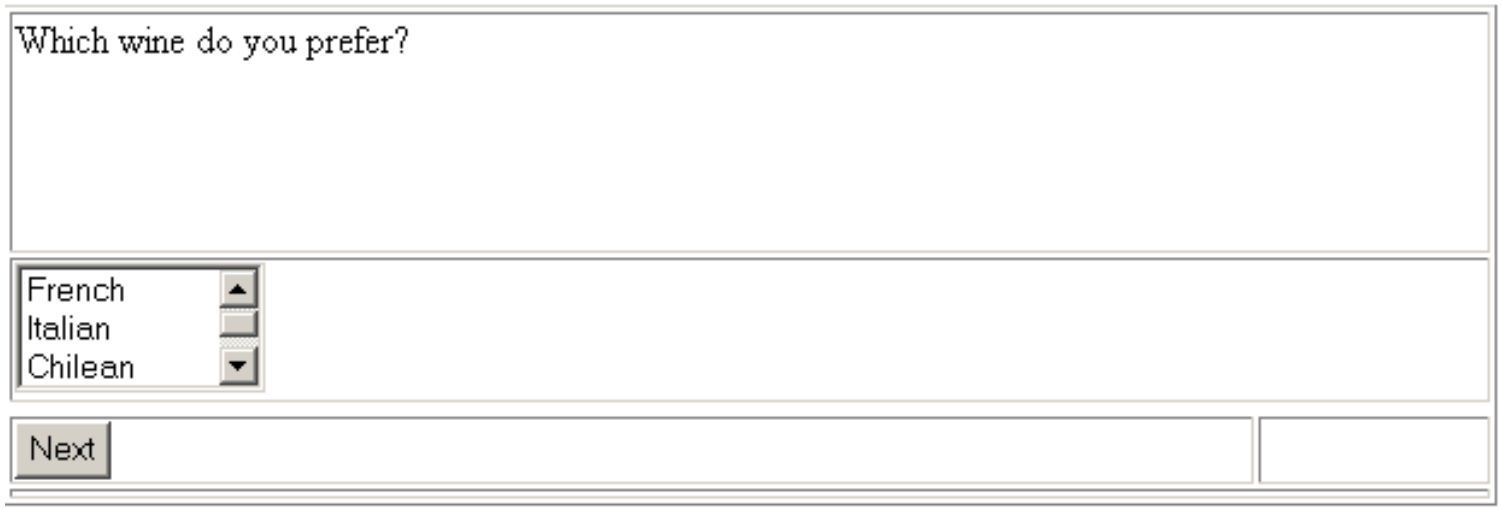
These effects should not be underestimated!

Picture	# of Meals	# of Shopping Trips	# of Overnight Trips	# of Sporting Events	# of Music Events
1. None	2.33 (.12)	3.30* (.16)	15.52 (.74)	1.57 (.19)	1.08 (.09)
2. Low freq.	2.16* (.12)	2.69* (.14)	15.12 (.73)	1.26* (.19)	1.02* (.09)
3. High freq.	2.61* (.13)	7.39* (.28)	15.06 (.75)	2.12* (.27)	1.37* (.11)
F *Df* *p*	3.47 2,1616 .03	166.88 2,1620 .00	.12 2,1604 .89	3.98 2,1622 .02	3.67 2,1621 .03

Instructions influence answers as well

- “Please count all shopping trips, from grocery store to department store”
- “Please count shopping trips to a department store only. Do not count grocery shopping”

Partial visibility affects answers



Drop-down menu with 3 out of 10 options initially displayed

Larger answer spaces elicit more information

- Increase number of words
- Increase number of themes
- Greater elaboration of answers (makes it easier to recode open answers)

Connections among answer spaces, labels, and symbols affect answers

- If you want people to give an answer of 2 digits, don't provide an answer box that is very wide
- Questions on dates
  - A MM/YYYY reduces errors from 55% to 4%

Guidelines for design choices

1. The size of the answer box should match the size of the desired answer
2. Make sure every answer option receives the same visual stimulus
3. Place ordinal scales consistently in a decremental or incremental order
4. Make sure that the visual midpoint of a scale coincides with the conceptual midpoint

Guidelines for design choices

5. If multiple items are presented on a single screen, be aware that correlations might be higher between items, especially when polar point (only endpoint labels) are used
6. Use fully labelled scales or polar point labels with numbers starting from 1 (unless your original questions used others)
7. Use a logical order of response options (e.g. a progression) and be aware that respondents extract meaning from that order

Guidelines for design choices

8. Use instructions right in front of the answer options and make sure respondents do not have to make an effort to remember them
9. Avoid using gratuitous visual language (e.g. pictures, numbers, colors)
10. When comparing results from different studies, make sure respondents get the same (visual) stimulus

Keep the design effects in mind when interpreting your results!

Elements that account for variance in survey responding:

1. (Size of) Answer space,
2. Spacing,
3. Ordering,
4. Color,
5. Numbers,
6. Pictures, sounds, and video's
7. Labels,
8. Symbols/Signals,
9. Instructions,
10. Visibility

Use these elements with care and compare them with other surveys!

Other issues:

- Progress indicator
- Routing
- Going back in the survey
- Error messages
- Don’t know and won’t tell
- Interactive features
- Randomizations
- Numbering
- Dependent interviewing
- Paradata