Survey – theoretical outline

GLCb1008 Introduction to Methodology of Social Sciences

Jiří Navrátil
jiri.navratil@fss.muni.cz

Lecture outline

- Context of survey design

- Sampling

- Data collection

- Asking questions

Explanation vs. understanding

- Idiographic (deterministic): understanding, revealing motives, reasons, justifications…
- Idiographic method – problem with zero
- Nomothetic methods (probabilistic): explanation, identifying causal relationships

Two logics of research process

(de Vaus 1996: 21)
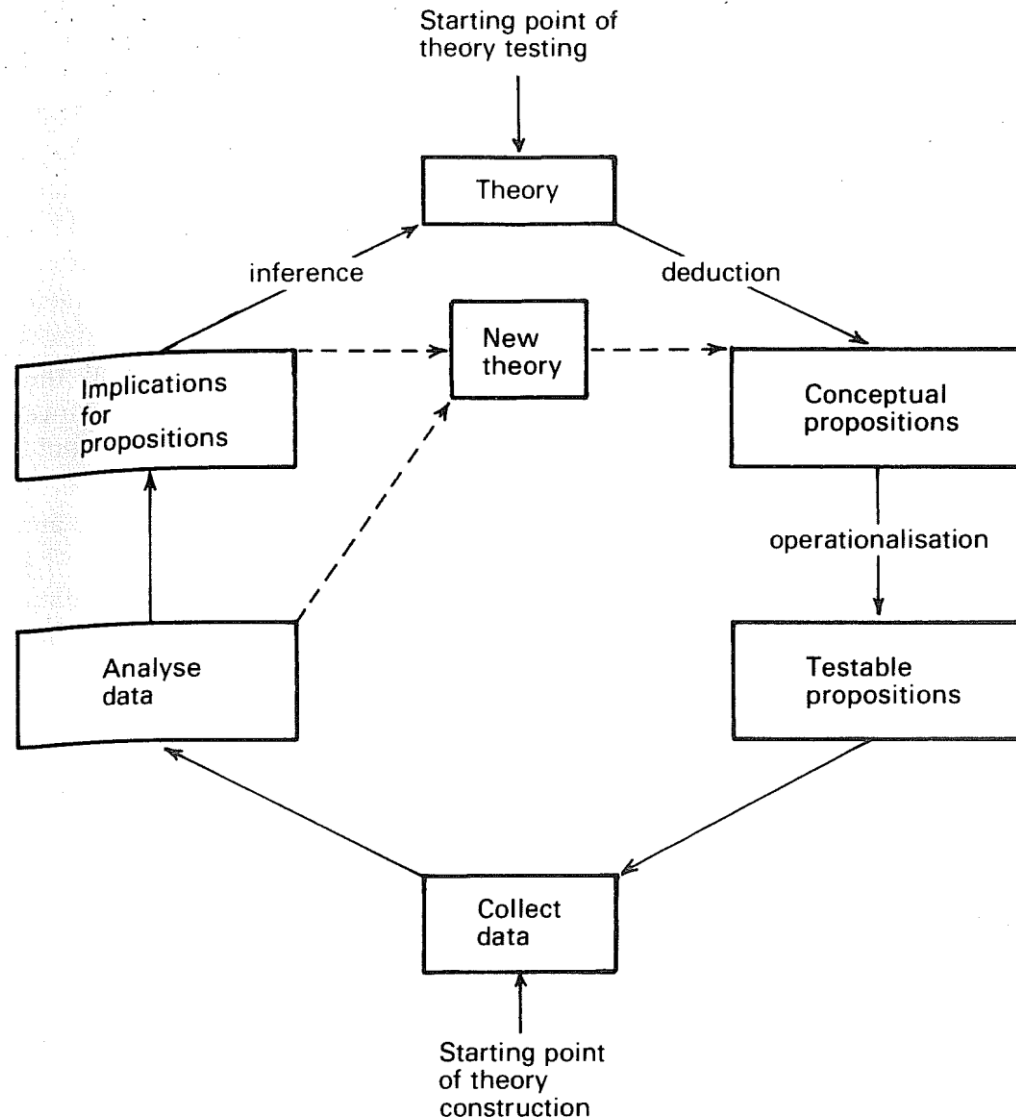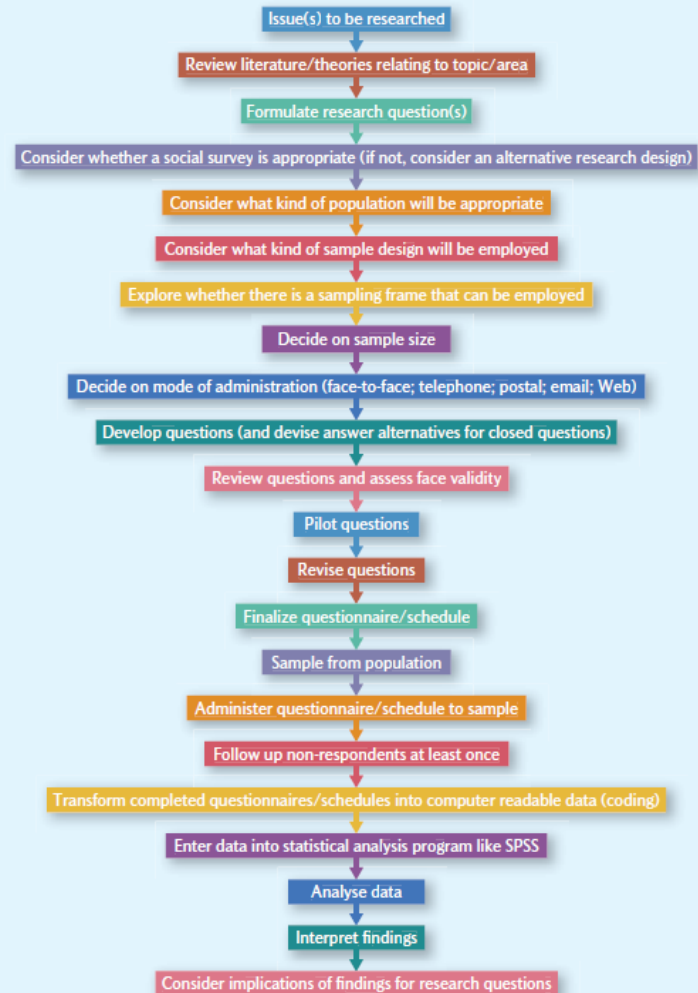
Main types of research strategies

- Experiment (small N)
- Case study (small N)
- Comparative design (small N)
- Longitudinal design (small/large N)
- Cross-sectional design (large N)

Social survey

- Probabilistic/nomothetic logic
- Deductive logic
- Cross-sectional design (large N)

Steps in conducting a social survey

- Issue(s) to be researched
- Review literature/theories relating to topic/area
- Formulate research question(s)
- Consider whether a social survey is appropriate (if not, consider an alternative research design)
- Consider what kind of population will be appropriate
- Consider what kind of sample design will be employed
- Explore whether there is a sampling frame that can be employed
- Decide on sample size
- Decide on mode of administration (face-to-face; telephone; postal; email; Web)
- Develop questions (and devise answer alternatives for closed questions)
- Review questions and assess face validity
- Pilot questions
- Revise questions
- Finalize questionnaire/schedule
- Sample from population
- Administer questionnaire/schedule to sample
- Follow up non-respondents at least once
- Transform completed questionnaires/schedules into computer readable data (coding)
- Enter data into statistical analysis program like SPSS
- Analyse data
- Interpret findings
- Consider implications of findings for research questions

Cross-sectional studies

- More than 1 case (in fact large N) in one time frame
- Collection of quantifiable data on the characteristics („variables") of these cases
- Logic of correlation between and among varibles
- Typically – social survey/questionnaire (but other data collection methods available too)
- Variables are not manipulated
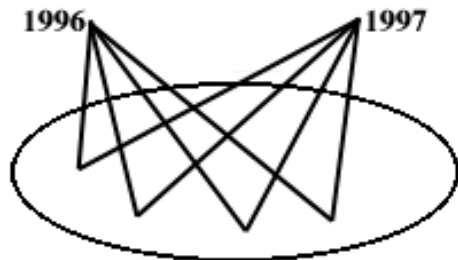
Time frame for cross-sectional studies

TREND

1996   1997

Trend Surveys draw respondents from the same population but the samples are different

COHORT

1996   1997

Cohort Surveys draw respondents from different populations like a grade 9 class in 1996 and another grade 9 class in 1997.

PANEL

1996   1997

Panel Surveys draw the same sample from the same population. the surveyor returns to the original respondents to follow up on their "growth" of attitudes to changing conditions.

EVS METHODOLOGY

Wave after wave, the European Values Study improved its methodological standards by increasing the harmonization of the procedures; applying the more recent developments in survey research; enhancing the transparency of the processes through monitoring and detailed documentation. In 2017 wave, the efforts to reach high level of data quality have been substantively increased. Read more about what has been done in EVS2017 to reach data quality in all the phases of the survey cycle.

The main features of the EVS Methodology are the following:

- Mode of data collection: F2F interview. In 2017 mixed-mode has been introduce as experiment
- Sample size: effective size=1000 in 1981; increased up to 1500 in 2008; 1200 in 2017
- Sampling: the use of quotes was admitted in the first waves, but since 2008 only probabilistic representative sample
- Target: Resident population older than 18 years old
- Translation and monitoring: centrally coordinated since 2008
- Data Access: data and documentation of all the waves are available free of charge. They are stored in the Data Archive for the Social Sciences of GESIS – Leibniz Institute in Cologne, Germany.

Sampling

- Population – all units (individuals, nations, organizations, cities …) from which we select sample and to which we refer with the results
- Can we handle whole population? (census …)
- Sample – part of units selected for analysis
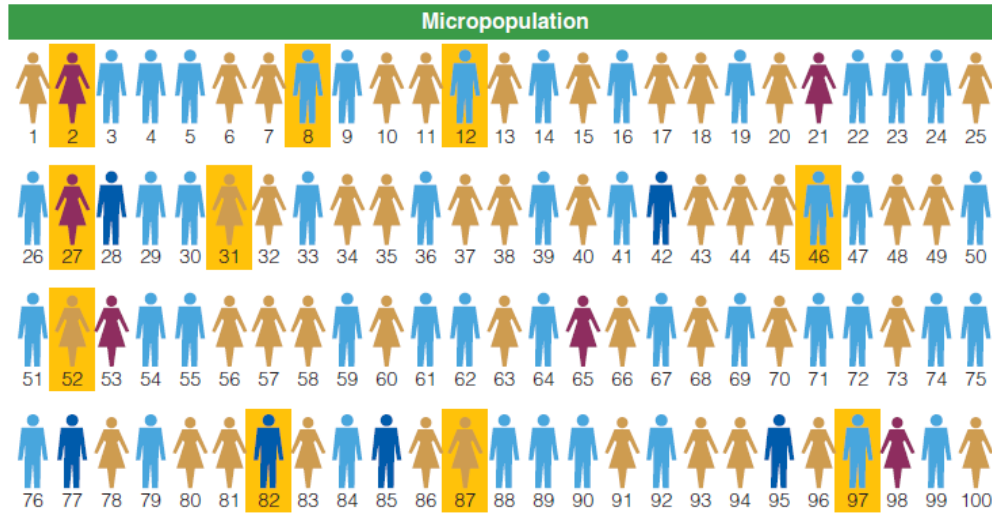- Probability sampling (logic of chance) vs. Non-probability (purpose)

Random sample

- Standard type of sampling in large N
- Representativity – known and unknown characteristics of population
- Option of statistical inference from the sample to the population
- Types of random sample: simple, systematic, stratified, multi-stage
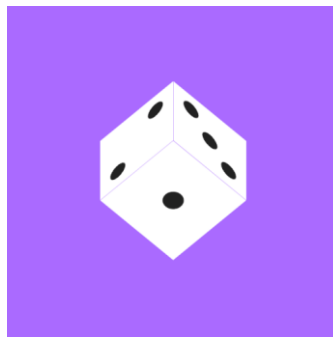
Simple random sample
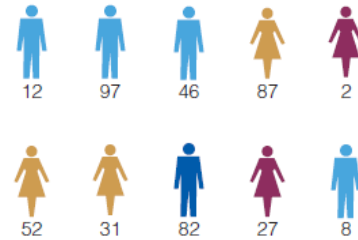
FIGURE 7-11

A Simple Random Sample. Having numbered everyone in the population, we can use a table of random numbers to select a representative sample from the overall population. Anyone whose number is chosen from the table is in the sample.

Systematic random sample
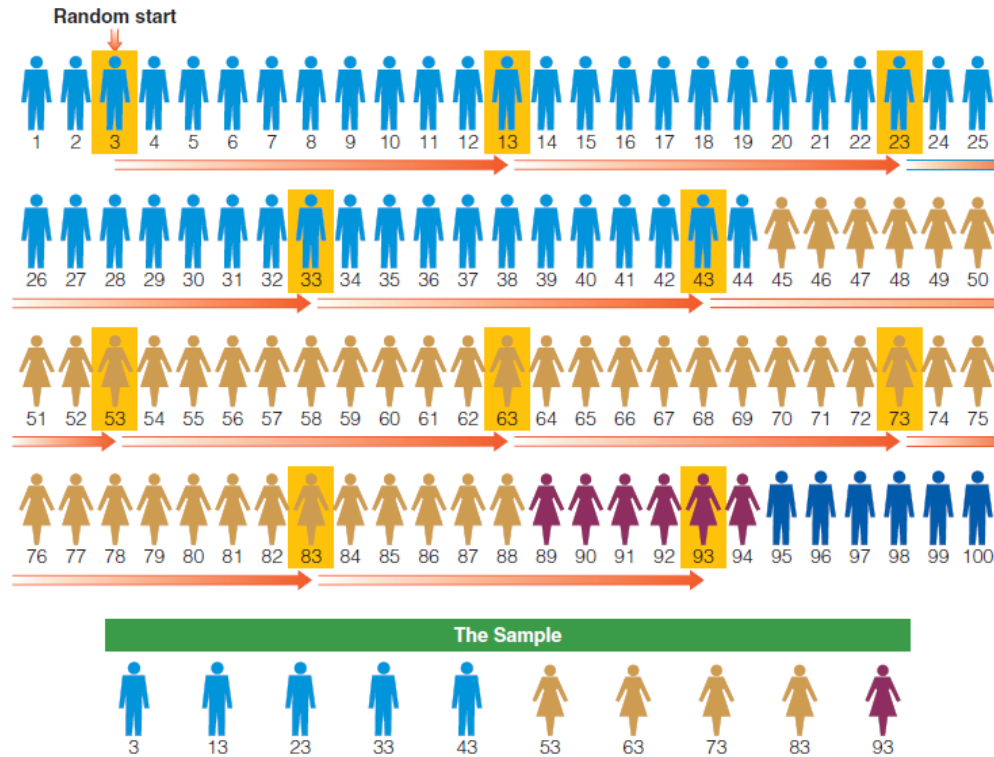
FIGURE 7-12

A Stratified, Systematic Sample with a Random Start. A stratified, systematic sample involves two stages. First the members of the population are gathered into homogeneous strata; this simple example merely uses gender and race as stratification variables, but more could be used. Then every *k*th (in this case, every tenth) person in the stratified arrangement is selected for the sample.

Multi-stage cluster sample

FIGURE 7-13

Multistage Cluster Sampling. In multistage cluster sampling, we begin by selecting a sample of the clusters (in this case, city blocks). Then, we make a list of the elements (households, in this case) and select a sample of elements from each of the selected clusters.

Non-probability sampling

- *convenience sampling* – who do you succeed to get in the sample (street distribution, relatives, friends, social media)

- *purposive sampling* – focusing at target group within population (study of student environmental activists – looking for the most visible ones at the faculty)

- *snowball sampling* – members of population are hard to reach (subcultural or social movement studies)
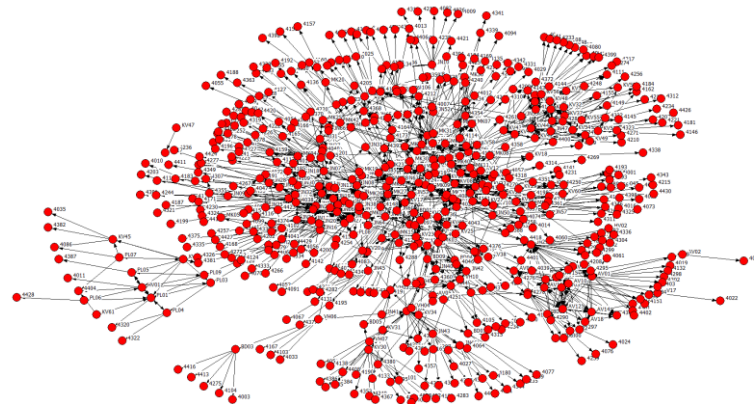
Non-probability sampling

- Quota sampling – when we know the characteristics of target population (typically: sex, education, ethnicity, age)

- Example: Czech adult population (attitudes towards the conflict in Ukraine)

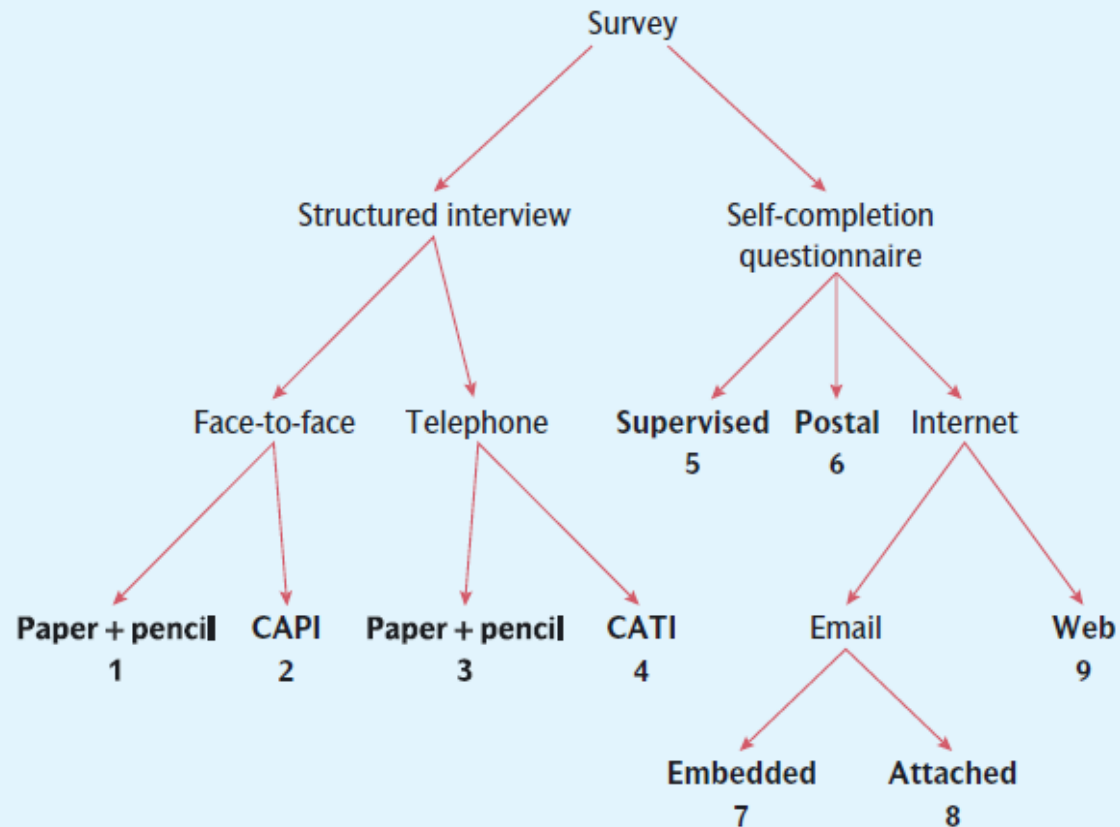| bydliště	kvóta
Hlavní město Praha	122
Středočeský kraj	129
Jihočeský kraj	61
Plzeňský kraj	55
Karlovarský kraj	27
Ústecký kraj	76
Liberecký kraj	42
Královéhradecký kraj	52
Pardubický kraj	49
Kraj Vysočina	48
Jihomoravský kraj	112
Olomoucký kraj	59
Zlínský kraj	55
Moravskoslezský kraj	113
pohlaví	kvóta
muž	490
žena	510
věk	kvóta
18 - 24	80
25 - 34	150
35 - 44	190
45 - 54	180
55 - 64	190
65 +	210
vzdělání	kvóta
ZŠ	136
SŠ bez maturity	326
SŠ s maturitou	351
VŠ	187

Example of quota sampling
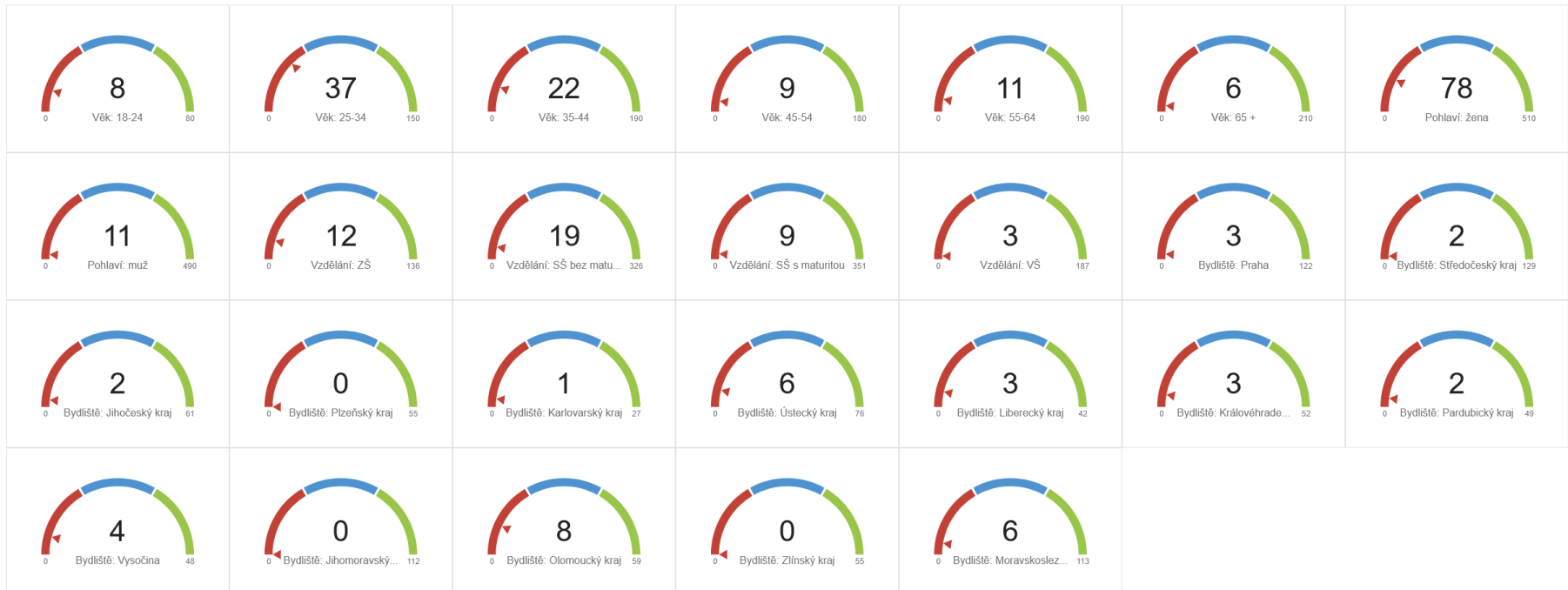
Sample size

- Absolute, not relative size of sample is important

- Increasing sample size – decreasing possibility of sampling error (random sampling)

- https://www.surveymonkey.com/mp/sample-size-calculator/

Types of data collection

Main modes of administration of a survey

Survey

- Structured interview
  - Face-to-face
    - Paper + pencil 1
    - CAPI 2
  - Telephone
    - Paper + pencil 3
    - CATI 4
- Self-completion questionnaire
  - Supervised 5
  - Postal 6
  - Internet
    - Email
      - Embedded 7
      - Attached 8
    - Web 9

Notes: CAPI is computer-assisted personal interviewing; CATI is computer-assisted telephone interviewing.

Data collection

1. Please tell us why you participated in this Pride parade? [01whyp]

....

....

....

2. In your opinion, who or what is to blame for discrimination against LGBT? [02whobl]

....

3. What should be done to address such discrimination? [03whatdo]

....

....

4. What do you think of the behavior of the police at the parade? [04police1-2]

	Not at all	Not very much	Somewhat	Quite	Very much
Cooperative	1❑	2❑	3❑	4❑	5❑
Aggressive	❑	❑	❑	❑	❑

5. And what about the other participants? Were they...? [05behdem1-2]

	Not at all	Not very much	Somewhat	Quite	Very much
Cheerful	1❑	2❑	3❑	4❑	5❑
Disorderly	❑	❑	❑	❑	❑

6. Were you at this parade (Tick as many as apply): [06comp1-8]

Alone?	1❑
With your partner?	1❑
With your children?	1❑
With friends?	1❑
With relatives?	1❑
With acquaintances?	1❑
With colleagues or fellow students?	1❑
With members of an organization you are a member of?	1❑

7. When did you make a firm decision to participate in the parade? [07decis]

1❑ The day of the demonstration 2❑ A few days before the demonstration 3❑ A few weeks before the demonstration 4❑ Over a month ago

Online tools

https://www.kobotoolbox.org/

https://www.surveymonkey.com/

Data matrix (column = variable; row = case)

Modes of survey administration compared

The strengths of email and Web-based surveys in relation to face-to-face interview, telephone interview, and postal questionnaire surveys

Issues to consider	Mode of survey administration				
	Face-to-face interview	Telephone interview	Postal questionnaire	Email	Web
Resource issues					
Is the cost of the mode of administration relatively low?	✓	✓✓	✓✓✓	✓✓✓	✓ (unless access to low-cost software)
Is the speed of the mode of administration relatively fast?	✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓
Is the cost of handling a dispersed sample relatively low?	✓ (✓✓ if clustered)	✓✓✓	✓✓✓	✓✓✓	✓✓✓
Does the researcher require little technical expertise for designing a questionnaire?	✓✓✓	✓✓✓	✓✓✓	✓✓	✓
Answering context issues					
Does the mode of administration give respondents the opportunity to consult others for information?	✓✓	✓	✓✓✓	✓✓✓	✓✓✓
Does the mode of administration minimize the impact of interviewers' characteristics (gender, class, ethnicity)?	✓	✓✓	✓✓✓	✓✓✓	✓✓✓
Does the mode of administration minimize the impact of the social desirability effect?	✓	✓✓	✓✓✓	✓✓✓	✓✓✓
Does the mode of administration allow control over the intrusion of others in answering questions?	✓✓✓	✓✓	✓	✓	✓
Does the mode of administration minimize the need for respondent to have certain skills to answer questions?	✓✓✓	✓✓✓	✓✓	✓ (because of the need to have online skills)	✓ (because of the need to have online skills)
Does the mode of administration enable respondents to be probed?	✓✓✓	✓✓✓	✓	✓✓	✓
Does the mode of administration reduce the likelihood of data entry errors by the researcher?	✓	✓	✓✓	✓	✓✓✓

Issues to consider	Mode of survey administration				
	Face-to-face interview	Telephone interview	Postal questionnaire	Email	Web
Sampling-related issues					
Does the mode of administration tend to produce a good response rate?	✓✓✓	✓✓	✓	✓	✓
Is the researcher able to control who responds (i.e. the person at whom it is targeted is the person who answers)?	✓✓✓	✓✓✓	✓✓	✓✓	✓✓
Is the mode of administration accessible to all sample members?	✓✓✓	✓✓	✓✓✓	✓ (because of the need for respondents to be accessible online)	✓ (because of the need for respondents to be accessible online)
Questionnaire issues					
Is the mode of administration suitable for long questionnaires?	✓✓✓	✓✓	✓✓	✓✓	✓✓
Is the mode of administration suitable for complex questions?	✓✓✓	✓	✓✓	✓✓	✓✓
Is the mode of administration suitable for open questions?	✓✓✓	✓✓	✓	✓✓	✓✓
Is the mode of administration suitable for filter questions?	✓✓✓ (especially if CAPI used)	✓✓✓ (especially if CATI used)	✓	✓	✓✓✓ (if allows jumping)
Does the mode of administration allow control over the order questions are answered?	✓✓✓	✓✓✓	✓	✓	✓✓
Is the mode of administration suitable for sensitive questions?	✓	✓✓	✓✓✓	✓✓✓	✓✓✓
Is the mode of administration less likely to result in non-response to some questions?	✓✓✓	✓✓✓	✓✓	✓✓	✓✓
Does the mode of administration allow the use of visual aids?	✓✓✓	✓	✓✓✓	✓✓	✓✓✓

Types of questions

- Open-ended vs. Closed questions (time, validity, reliability, comparability, clarification, easiness, spontaneity, exhaustiveness)
- Personal factual, others factual, informant factual, attitudes, beliefs, normative values, knowledge

How to design the questions

- Follow research question!
- Be specific! (What is your opinion about contemporary party politics?)
- Avoid ambiguous questions! (How often do you eat fast food?)
- Avoid long questions! (On a scale of 1-10, how likely are you to recommend political party you voted for in last elections to a friend or colleague who works in a similar industry as you do?)
- Avoid double-barrelled questions! (Do you think that students should have more classes about history and culture?)
- Avoid leading questions! (Do you think that our country is a true democracy?)
- Think about filters and conditions! (What movie did you see in a cinema last month?)
- Avoid using jargon and technical terms! (What is your opinion of the efficacy of the Twitter´s new search algorithm?)
- Think about the respondent´s capacities, memory and knowledge! (When did you drink alcohol for the first time?)
- Always do the piloting!
- Use existing questions!

References

Bryman, Alan. 2012. *Social Research Methods*. Oxford: Oxford University Press.

Blaikie, Norman. 2000. *Designing social research: the logic of anticipation*. Cambridge: Polity Press.

Biemer, Paul P., Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz and Seymour Sudman (eds.). 1991. *Measurement Errors in Surveys*. New York: John Wiley & Sons.

De Vaus, David A. 1996. *Surveys in Social Research*. London: University College London Press Limited.

Dillman, Don A. 2007. *Mail and Internet Surveys. The Tailored Design Method.* New York: John Wiley & Sons.

Fowler, Floyd J. 1995. *Improving Survey Questions: Design and Evaluation*. Thousand Oaks: Sage Publications.

Groves, Robert M., Floyd J. Fowler, Mick P. Couper, James M. Lepkowski, Eleanor Singer and Roger Tourangeau. 2004. *Survey Methodology*. New York: John Wiley & Sons.