

# Quality Issues in Survey Research

DANS Symposium on  
Survey Data Quality  
Amsterdam, April 9, 2008

Joop J. Hox & Edith D. de Leeuw, Utrecht University

---

---

---

---


---

---

---

---

## What is Quality?

- W. Edwards Deming
  - Statistician & quality control expert
  - WW II: Statistical quality control
  - Influential in Japan's reconstruction after WW II
  - Forgotten in US, picked up around 1980 by Joseph Juran
- Deming: build quality into the process, do not rely on quality checks
  - Focus on quality, costs will then decrease
  - Focus on costs, quality will decrease but costs will not

*Quality - conditions make it possible  
Total costs*

2

---

---

---

---


---

---

---

---

## What is Quality?

- Juran: Quality is "Fitness for Use"
- Juran started thinking about quality in a wider context
  - Attention to process quality
  - Collect process data to monitor quality
  - Management tools to maintain quality
  - Maintaining quality as a continuous process
- This has strongly influenced the ISO (International standard setting)

3

---

---

---

---

---

---

---

---

## What is Quality?

- Juran: Quality is "Fitness for Use"
- Multi-dimensional, subjective notion depending on product, production process, and user
  - Construction quality (e.g., IKEA quality)
  - Operational quality (e.g., max. # of defects)
  - Sampling statistics quality (e.g., Mean Square Error)
  - Total Quality Management (e.g., Process quality)



4

---

---

---

---

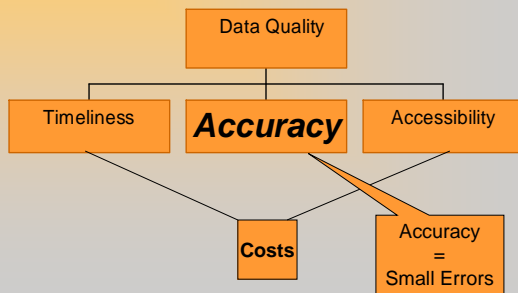
---

---

---

---

## Quality in Surveys: Framework



5

---

---

---

---

---

---

---

---

## Total Survey Quality

- More than just minimizing total survey error
- Add fitness for use
  - Process quality, timeliness, utility
- Wider quality definition taken up in survey methodology 1990 ...
  - ASA survey conferences on survey measurement and process quality
  - Various International workshops
  - Researchers from official statistics & academia

6

---

---

---

---

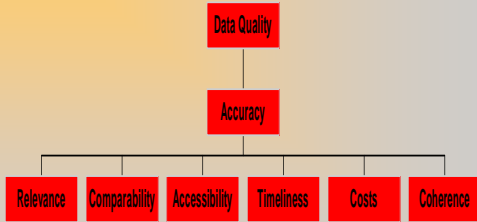
---

---

---

---

## Eurostat Quality Framework



7

---

---

---

---

---

---

---

---

## Survey Quality and the Survey Industry

- ISO 20252: Market, opinion and social research – Vocabulary and service requirements, requires:
  - Description of sampling procedure
  - Translated questionnaires independently checked by a second translator
  - A pretest of *all* self-administered questionnaires
  - Observing and checking data collection
  - Validation of data collection (e.g. call-back)
- NEN-ISO 20252 is about 80 pages...

8

---

---

---

---

---

---

---

---

## Survey Quality and Access Panels

- ISO 26362 (draft): Access panels in market, opinion and social research – Vocabulary and service requirements, requires:
  - Where applicable (e.g., quality management system): ISO 20252 compliant
  - Recruitment:
    - panel members recruited from a documented source
    - confirmation of identity and provided initial profile data
    - yearly update of member profiles
  - Cookies: if used explain in general what they are, and the differences between cookies for routing and *cookies for capturing metadata*
  - Data validation
    - completion time, # of unanswered questions
    - Comparison of at least 1 question with profile data

9

---

---

---

---

---

---

---

---

## Professional Organizations' View on Quality Standards

- AAPOR: Code of Ethics, Best Practices, Standard Rules for Response Rates
- ESOMAR: Similar
  - AAPOR and ESOMAR have been influenced by and have also influenced the ISO standards
- AAPOR on access panels: "Reporting a margin of error or sampling error for ... with opt-in ... samples ... is misleading and connotes a validity that may not be present."

10

---

---

---

---

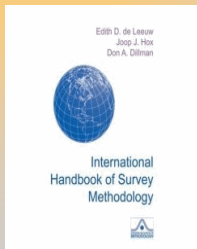
---

---

---

---

## Survey Methodologists and Statisticians on Quality



11

---

---

---

---

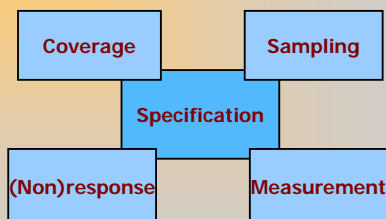
---

---

---

---

## Corner Stones of Data Quality



12

---

---

---

---

---

---

---

---

## Specification Error

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

13

---

---

---

---

---

---

---

---

## Specification Error

- Example 1 (Biemer & Lyberg, 2003):
  - Intended concept is 'value of a parcel of land if it were sold on a fair market today'
    - Q1: "For what price would you sell this parcel of land?"
    - Q2: "What do you think is the current market value of this piece of land? "

14

---

---

---

---

---

---

---

---

## Coverage Error

- Sampling frame must include all units of population of interest (once), if not:
- Frame Coverage Errors
  - Non-sampling error
- Errors arising from construction of sampling frame
  - Omissions
  - Erroneous inclusions
  - Duplications

15

---

---

---

---

---

---

---

---

## *Sampling Error*

- In principle well understood
- BUT
  - Complications due to weighting, stratification
  - Addressed in detail by official statistics
  - Largely ignored by social scientists

16

---

---

---

---

---

---

---

---

## *Nonresponse Error*

- Severe problems due to nonresponse & panel attrition
- Nonresponse leads to bias if nontrivial + related to variable under study
  - Difficult to prevent
  - Difficult to cope with

17

---

---

---

---

---

---

---

---

## *Measurement Error*

- A respondent's answer to a question is inaccurate, departs from the 'true' value
- Measurement errors are associated with the data collection process itself
- There are three main sources of measurement error:
  - **Questionnaire**
  - **Respondent**
  - **Method (mode) of data collection**
    - When **interviewers** are used for data collection, the interviewer is a fourth source of error

18

---

---

---

---

---

---

---

---

## What Should We Do?

- Use of new technology (mobile phone, web) is unavoidable
  - 'Resistance is futile'
- Better theories on the question-response process, using modern cognitive psychology & combine with research on human-computer interface
- New statistics on 'amount of representativeness': R-indexes, Bayesian statistics?

19

---

---

---

---

---

---

---

---

## Suggested Readings

- Edith de Leeuw, Joop Hox, Don Dillman (eds) 2008. *International Handbook of Survey Methodology*. Taylor and Francis, Psychology Press.
  - [www.xs4all.nl/~EdithL/surveyhandbook](http://www.xs4all.nl/~EdithL/surveyhandbook)
- Paul Biemer & Lars Lyberg (2003). *Introduction to Survey Quality*. Wiley.

20

---

---

---

---

---

---

---

---