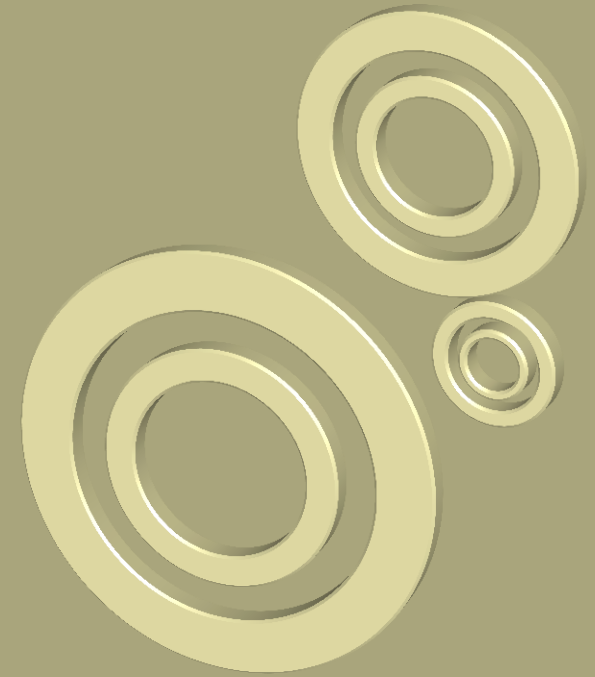


Introduction to Issues on Quality

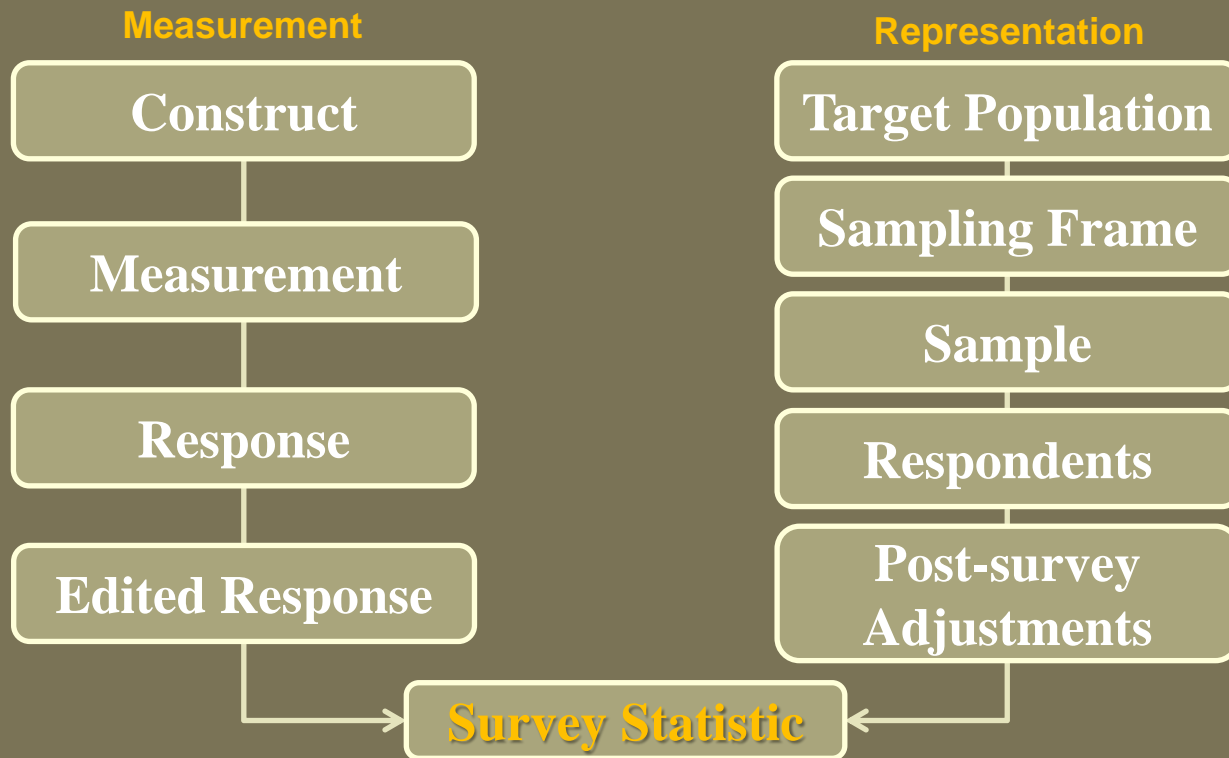


JULY 2011

Afsaneh Yazdani

Survey Quality

Life Cycle of a Survey from a Design Perspective



Survey Quality

What can affect **representativeness** of a sample?

- Sampling Design
- Nonresponse
- Measurement Problems
 - Inability to recall answers to questions
 - Leading questions
 - Unclear wording of questions, ...



Survey Quality

ERROR

Term error does not imply mistakes , instead, it refers to deviation of what is desired in the survey process from what is attained.



Survey Quality

ERROR

Term error does not imply mistakes , instead, it refers to deviation of what is desired in the survey process from what is attained.

The job of a survey designer is to minimize error in survey statistics by making design and estimation choices that minimize the **gap between two successive stages**

Survey Quality

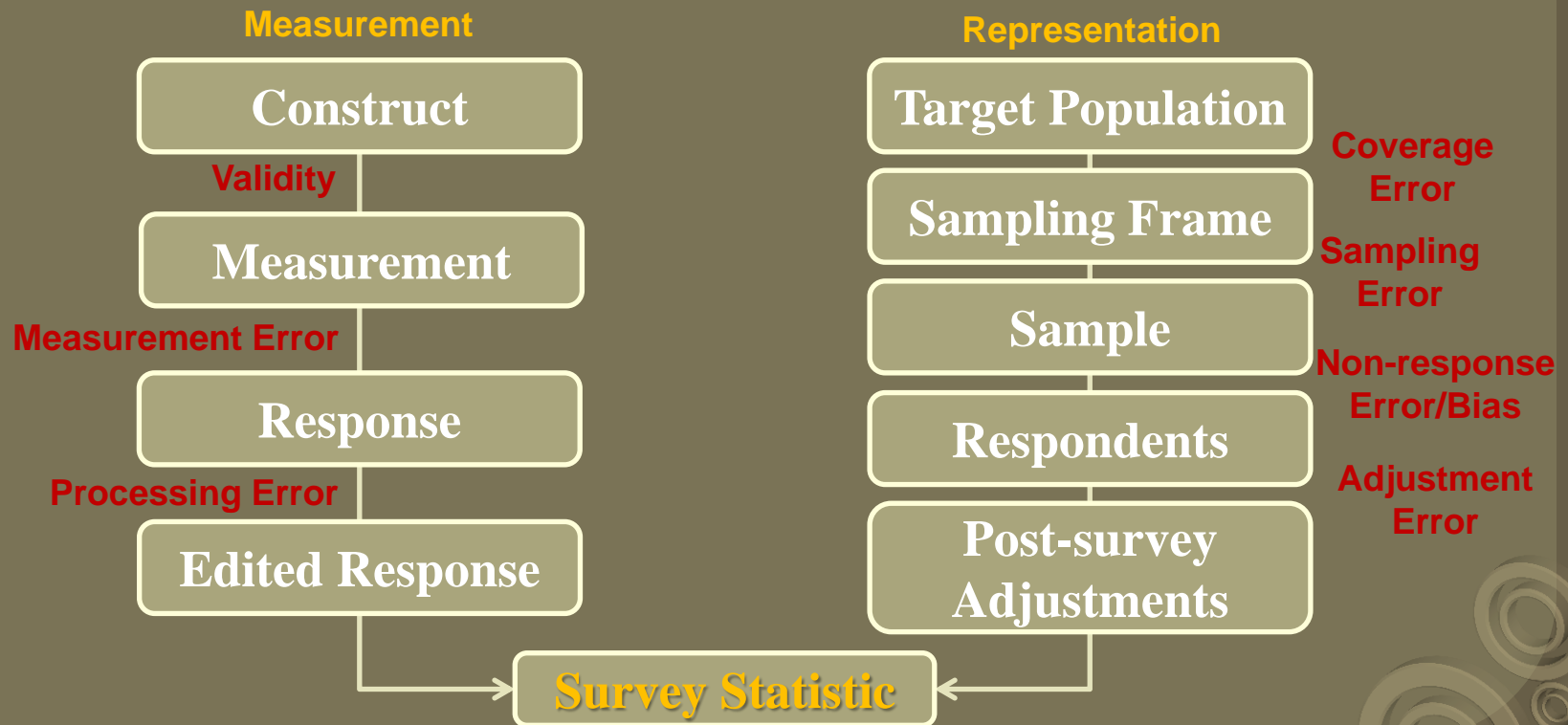
ERROR

Term error does not imply mistakes , instead, it refers to deviation of what is desired in the survey process from what is attained.

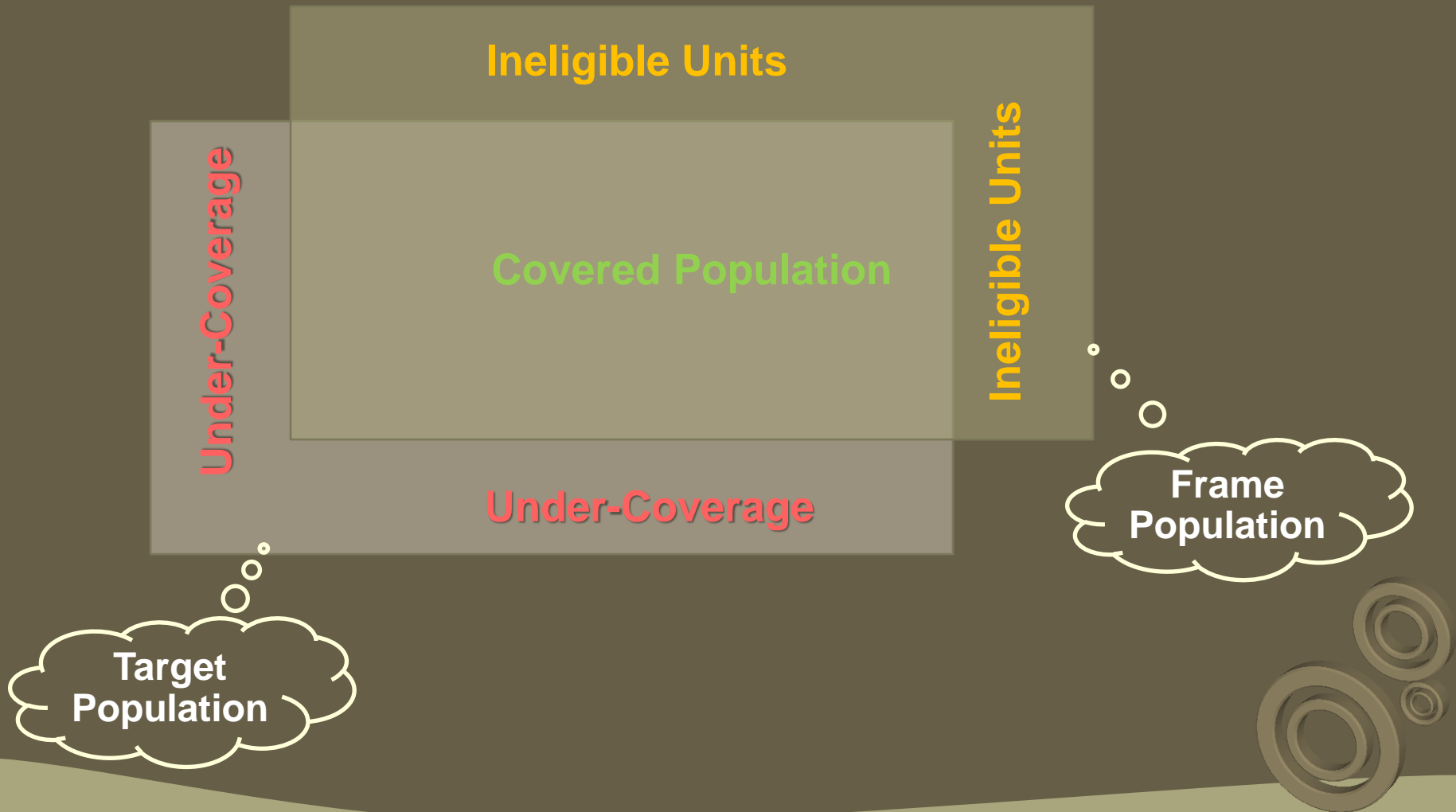
In other words, minimizing the total survey error

Survey Quality

Life Cycle of a Survey from a Quality Perspective



Survey Quality



Survey Quality

Other Notions of Survey Quality

Fitness for use: Acknowledging that different users of the same estimate may have different purposes:

- Credibility
- Relevance
- Timeliness



Survey Quality

Other Notions of Survey Quality

Fitness for use: Acknowledging that different users of the same estimate may have different purposes:

- Credibility
- Relevance
- Timeliness

The extent to which the producer of the information is judged by the user to be free of any particular point of view

Survey Quality

Other Notions of Survey Quality

Fitness for use: Acknowledging that different users of the same estimate may have different purposes:

- Credibility
- Relevance
- Timeliness

Measuring a construct quite similar in meaning to the user's main concerns. Providing the information demanded.

Survey Quality

Other Notions of Survey Quality

Fitness for use: Acknowledging that different users of the same estimate may have different purposes:

- Credibility
- Relevance
- Timeliness

Availability at the time it is needed for decisions based on the information

Survey Quality: Coverage Error

Coverage Properties

- Under-Coverage
- Ineligible Units
- Duplication



Survey Quality: Coverage Error

Coverage Properties

- Under-Coverage
- Ineligible Units
- Duplication

A result of:

- Processes used to construct the sampling frame
- Frame being out-of-date

Survey Quality: Coverage Error

Coverage Properties

- Under-Coverage
- Ineligible Units
 - Less difficult than Under-coverage
 - Better to screening before the survey
 - Screening during the survey causes reduction in sample size
- Duplication

Survey Quality: Coverage Error

Coverage Properties

- Under-Coverage
- Ineligible Units
- Duplication
 - A single target population element is associated with multiple units in frame
 - Duplicates have more chance of selection and can be over-represented.
 - Can be detected at the time of selection

Survey Quality: Coverage Error

Remedies for Coverage Errors

- Under-Coverage
- Ineligible Units
- Duplication



Half-Open Interval:

If within the interval of two units in the list, there are any newly constructed or missed units, it is added to the list, and interviewed. All the addresses in half-open interval have the same probability of selection, and new are added to the frame.

Survey Quality: Coverage Error

Remedies for Coverage Errors

- Under-Coverage
- Ineligible Units
- Duplication



Multiple Frame Designs:

Supplementing a principal frame that provides nearly complete coverage of the target population by a frame that provides better or unique coverage of absent units.

Survey Quality: Coverage Error

Remedies for Coverage Errors

- Under-Coverage
- Ineligible Units
- Duplication



- **Multiplicity Sampling:** Making network of units such as Snowball sampling

Survey Quality: Coverage Error

Remedies for Coverage Errors

- Under-Coverage
- Ineligible Units
- Duplication

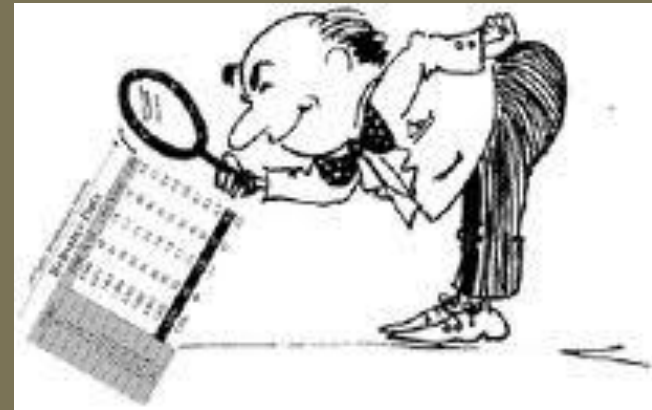


- Eliminating the duplicates or ineligible units
- Adjusting the weight for compensating the coverage errors

Survey Quality: Measurement Error

Sources of Measurement Error

- Interviewer
- Measurement Instrument
- Respondent



Survey Quality: Measurement Error

Sources of Measurement Error

- Interviewer
- Measurement Instrument
- Respondent

Interviewer
(if carefully selected, well-trained, highly motivated and well supervised and monitored) can attract cooperation, can be effective in locating and identifying eligible units, can assist in clarifying, probing, and motivating respondent to provide complete and accurate response

Survey Quality: Measurement Error

Sources of Measurement Error

- Interviewer
- Measurement Instrument
- Respondent

Presence of even a good Interviewer can be **harmful** for **sensitive** questions

Interviewer
(if carefully selected, well-trained, highly motivated and well supervised and monitored) can attract cooperation, can be effective in locating and identifying eligible units, can assist in clarifying, probing, and motivating respondent to provide complete and accurate response

Survey Quality: Measurement Error

Sources of Measurement Error

- Interviewer
- Measurement Instrument
- Respondent

A Questionnaire,
or another
instrument

Should be well-
designed, well-
stated, standard...

Survey Quality: Measurement Error

Sources of Measurement Error

- Interviewer
- Measurement Instrument
- Respondent

Unable to comprehend
the questions

Difficulty in remembering
the answers

Privacy

Survey Quality: Non-Response Error

Non-Response Error

Is **failure** to obtain measurements on **eligible** sampled units.

- Unit Non-response
- Item Non-response



Survey Quality: Non-Response Error

Non-Response Error

Is **failure** to obtain measurements on **eligible** sampled units.

- Unit Non-response
- Item Non-response

Failure is Complete

Survey Quality: Non-Response Error

Non-Response Error

Is **failure** to obtain measurements on **eligible** sampled units.

- Unit Non-response
- Item Non-response

Failure is Partial

Survey Quality: Non-Response Error

Non-Response Error

Is **failure** to obtain measurements on **eligible** sampled units.

- Unit Non-response
- Item Non-response

Can affect the quality of survey estimate if respondents and non-respondents differ

Survey Quality: Non-Response Error

Types of Unit Non-Response

- Non-Contact or failure to deliver the survey request
- Refusal to participate
- Inability to participate



Survey Quality: Non-Response Error

Remedies for Unit Non-Response

- Increase number of call, change timing of the calls, using lists with better addresses to help locating the unit
- Convince the sample unit about the confidentiality; well-know institutes conducting the survey; persuasion letter; two-phase sampling
- Interviewer language skills, changing the method of data collection



Survey Quality: Non-Response Error

Remedies for Item Non-Response

- Change the wording of a questionnaire
- Special training on the item
- Use suitable indirect techniques for sensitive questions



Survey Quality: Non-Response Error

Compensating for Unit Non-response

Weighting to adjust for unit non-response

Compensating for Item Non-response

Imputation: Placement of one or more estimated answers into a field of a data record that previously had no data or had incorrect or implausible data.



Survey Quality: Non-Response Error

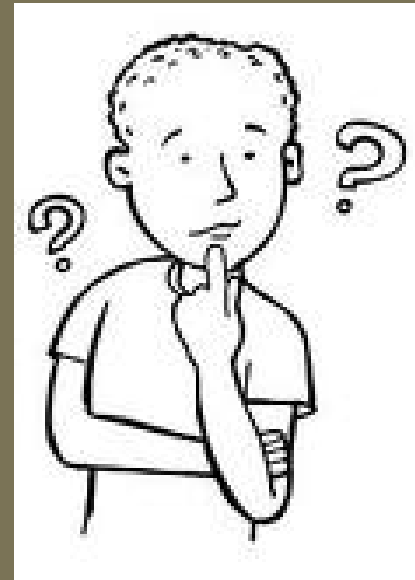
Impacts of Imputation

- **Advantage:** always the same number of cases are in analysis; procedure used for imputation is known
- **Disadvantages:** Imputed value seems equivalent to “made-up data”, leads to underestimation of standard errors of estimates.

Survey Quality: Non-Response Error

Imputation Methods

- Mean Value Imputation
- Regression Imputation
- Hot-deck Imputation
- Multiple Imputation



Survey Quality: Non-Response Error

Imputation Methods: Mean Value Imputation

As the simplest method, replaces the missing values for a particular variable by mean of the item respondent values.

Disadvantage

The distribution of the values in the imputed data is distorted.



Survey Quality: Non-Response Error

Imputation Methods: Mean Value Imputation

As the simplest method, replace the missing values for a particular variable by mean of the item respondent values.

Disadvantage

The distribution of the values in the imputed data is distorted.

**Better to be done in
sub-groups**

Survey Quality: Non-Response Error

Imputation Methods: Regression Imputation

Using a regression model to predict a value for the missing value.

May require that variables be imputed in sequence so that predictors in model have their imputed values themselves.



Survey Quality: Non-Response Error

Imputation Methods: Hot-deck

Sorts the data over the variables which are correlated to the variable with missing value

If for the first record there is a missing value, the mean of the total sample or the mean of the sub-group is replaced. (initial imputed value is called cold-deck value)

If not, the reported value for it, is stored as “hot-deck value” then the first missing value is replace with the hot-deck value.



Survey Quality: Non-Response Error

Imputation Methods: Hot-deck



Appropriate for large-scale imputation processes.

Survey Quality: Non-Response Error

Imputation Methods: Multiple Imputation

Creates multiple imputed datasets, each one based on a different realization of an imputation model for each item imputed.

Variation in estimates across these multiple data sets permits estimation of overall variation, including both sampling and imputation variance



Survey Quality: Deterministic Tests

Few Deterministic Tests to help us edit data

Range Test

- Range of admissible values

If-then test

- If-then relation Between variables

Ratio Control Test

- The lower and upper bound for the ratio

Zero Control Test

- Checking the sum with the element, like in balance tables

Internal Consistency

- Checking the internal consistency of variables

Probabilistic Test

- Checking for outliers