

SURVEY METHODS and SAMPLING



Cenker Burak METİN

Head of Survey and Questionnaire Design Group

Purpose of the course

The course will provide an overview of survey methods. It covers many issues about basic survey concepts useful to all participants. Some subjects included in the course are more technical than others. As a result, the course aims to

- Have an idea about survey processes
- Provide proficiency in sampling and estimation

Expectations

- After the course, participants are expected to have knowledge of how to manage the survey process.
- It is also expected that they will determine the processes that need to be developed in their department.
- For those who work on sampling, they are expected to increase the quality of estimation.

Question Form

(Rate from 1- 5, with 1 as the least and 5 as highest)

How do you describe yourself about your **sampling** knowledge level?

5	4	3	2	1
☐	☐	☐	☐	☐
☐	☐	☐	☐	☐

How do you describe yourself about your **R** knowledge level?

What is a Survey?

- Definition of a statistical survey and basic concepts
- Survey planning and steps of a survey
- Errors in surveys

Cenker Burak METİN

Head of Survey and Questionnaire Design Group



Definition of a statistical survey and basic concepts

What is a survey?

Any activity that collects information in an organised and methodical manner about characteristics of interest from some or all units of a population using well-defined concepts, methods and procedures, and compiles such information into a useful summary form (Survey Methods and Practices, 2010)

The term survey covers any activity that collects or acquires statistical data. Included are censuses, sample surveys, the collection of data from administrative records and derived statistical activities.

Census

A census is a survey conducted on all observation objects belonging to a particular population or universe

Pros of a CENSUS

- No sampling error
- Suitable for benchmark data
- Detailed information about small sub-groups



Census in Bethlehem (Pieter Bruegel, 1566)

Cons of a CENSUS

- Difficult to enumerate all units
- Higher costs
- Generally takes longer

Sample Survey

A sample survey is a survey which is carried out using a sampling method, i.e. in which a portion only, and not the whole population is surveyed

Pros of a SAMPLE

- Lower costs
- Less time
- The results can be very representative of the actual population

Cons of a SAMPLE

- Not be representative of the total population
- Not suitable for producing benchmark data
- 'Sampling' error
- Not detailed information about small sub-groups

Administrative Data Systems

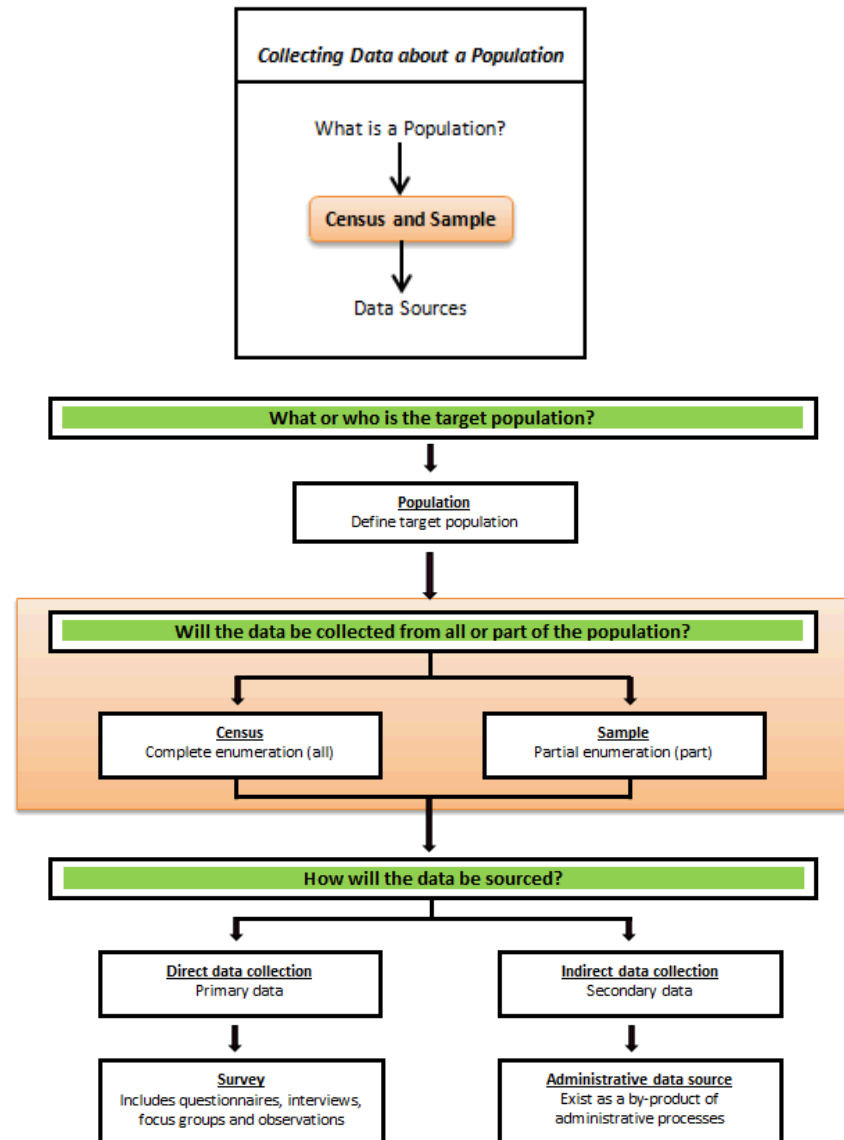
Administrative data collection is the set of activities involved in the collection, processing, storage and dissemination of statistical data from one or more administrative sources.

- Have a huge advantage in data collection costs and respondent burden
- Concepts and definitions must be carefully assessed for administrative program
- Lack of control over the quality of the data
- Processing work to be done
- Privacy concerns

The main statistical uses of administrative data are:

- Direct tabulation or analysis
- Indirect estimation
- Survey frames
- Survey evaluation





Survey planning and steps of a survey

1. Formulation of the Statement of Objectives
2. Selection of a survey frame
3. Determination of the sample design
4. Questionnaire design
5. Data collection
6. Data capture and coding
7. Editing and imputation
8. Estimation
9. Data analysis
10. Data dissemination
11. Documentation

(Source: Survey Methods and Practices, 2010)

1. Formulation of the Statement of Objectives

what is to be included in the survey?

what is to be excluded?

what the client needs to know? vs. what would be nice to know?

The steps in the process:

a) The Information Needs (State the Problem)

What are the overall information needs of the survey?

b) The Users and Uses of the Data

Who will use the data and how will they use it?

c) Concepts and Operational Definitions

What definitions will be used by the survey?

d) Survey Content

What are the specific topic areas to be covered by the survey?

e) The Analysis Plan (Proposed Tabulations)

Has an analysis plan with proposed tabulations been prepared?

2. Selection of a Survey Frame

The survey frame provides the means of identifying and contacting the units of the survey population.

The frame is in the form of a list, for example:

- a physical list such as a data file, computer printout or a telephone book;
- a conceptual list, for example a list of all vehicles that enter the parking lot of a shopping centre between 9:00 a.m. and 8:00 p.m. on any given day;
- a geographic list in which the units on the list correspond to geographical areas and the units within the geographical areas are households, farms, businesses, etc.



3. Determination of the Sample Design

A sample that has the ability to represent the population should be

- sufficient in size,
- similar to the distribution in the population in terms of variety and ratio,
- chosen with one of the probability sampling methods,
- no bias in its selection.

4. Questionnaire design

Questionnaire (or form) is a group or sequence of questions designed to obtain information on a subject from a respondent.

A well-designed questionnaire should:

- collect data efficiently with a minimum number of errors and inconsistencies;
- be respondent friendly and interviewer friendly (if interviewer-assisted);
- lead to an overall reduction in the cost and time associated with data collection.

5. Data collection

- ✓ Should the questionnaire be administered by an interviewer?
- ✓ Should a combination of methods be used?
- ✓ Should the questionnaire be paper or computer-based?
- ✓ Should administrative data be used to collect some of the survey data?
- ✓ Should data collection for several surveys be combined?

The basic methods of data collection are:

- Self-Enumeration (PAPI, CASI)
- Interviewer-assisted (PAPI, CAPI, CATI)

Mode of administration	Data capture instrument	
	Paper questionnaire	Electronic questionnaire
Self-completion	Self-administered questionnaire, diaries	Internet questionnaire, Computer Assisted Self Interview (CASI)
Interview	Interviewer-administered questionnaire	Computer Assisted Personal Interview (CAPI), Computer Assisted Telephone Interview (CATI)

Source: Quality Guidelines for Official Statistics (2002)

6. Data Capture and Coding

Data capture is the process of transferring respondents information onto a computerized system.

- Data was captured from forms by physically keying it into a computer in paper based collection method. However, the majority of data capture now occurs via an electronic process.
- With electronic capture of the information, data entry allows the collectors of the information to make queries at the source and alert the operator to possible errors in the data entry process.

Coding is the process of assigning a numerical value to responses to facilitate data capture and processing in general.

- When determining the coding scheme, the goal should be to classify responses into a meaningful set of exhaustive and mutually exclusive categories that bring out the essential pattern of responses.

7. Editing and Imputation

Editing is the application of checks to identify missing, invalid or inconsistent entries that point to data records that are potentially in error.

Edits can range from simple manual checks performed by interviewers in the field to complex verifications performed by a computer program.

- editing should be performed at several stages of the survey.
- edits applied at each stage should not contradict edits at some other stage
- editing should be used to provide information about the survey process, either in the form of quality measures for the current survey or to suggest improvements for future surveys.
- corrections should not directly or indirectly cause other errors,
- autocorrect should be used whenever possible.



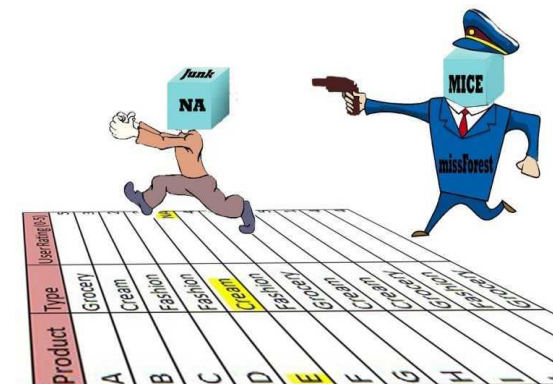
7.Editing and Imputation

Imputation is a process used to determine and assign replacement values to resolve problems of missing, invalid or inconsistent data.

Although imputation can improve the quality of the final data, care should be taken in order to choose an appropriate imputation methodology.

The following are some guidelines for imputation:

- Imputed records should closely resemble the failed edit record.
- The imputation methods should be chosen carefully, considering the type of data to be imputed.
- The imputation system should be able to handle any pattern of missing or inconsistent fields.
- The imputation system should be thought out, specified, programmed and tested in advance.
- The process should be automated, objective, reproducible and efficient.
- The imputation method should aim to reduce the nonresponse bias



8. Estimation

- The principle behind estimation is that each sample unit represents not only itself, but also several units of the survey population. Determining the weight is an important part of the estimation process.
- The principal purpose of weighting is to obtain as accurate parameter estimates as possible with the chosen sampling and estimation procedures.
- The basic form of weight (called sampling weight or design weight) is defined as the inverse of the inclusion probability of a selected element.
 - ✓ Estimation must take into account the sample design.
 - ✓ The initial design weights should be adjusted for nonresponse.
 - ✓ Auxiliary information should be used
 - ✓ The sample design and sample allocation should be used
 - ✓ Outliers can lead to large sampling variability in the estimates
 - ✓ Estimation of sampling error

9. Data analysis

- Data analysis involves summarizing the data and interpreting their meaning in a way that provides clear answers to questions that initiated the survey.
- Often, it consists of examining tables, charts and various summary measures, such as frequency distributions and averages to summarize the data.
- Statistical methods for analysing data can be characterised according to the type of data to which they are applied.
- The field of survey statistics usually deals with cross-sectional data describing many different individuals or units at a single point in time.
- Time series data describing a single entity across time.
- Longitudinal data blends characteristics of both crosssectional and time series data.

10. Data dissemination

Data dissemination is the release of the survey data to users through various media. When delivering the results to users, it is important to ensure that the information is accurate, complete, accessible, understandable, usable, timely, meets confidentiality requirements and is appropriately priced.

As a general rule the report should include the following:

- **Executive summary** (summarizing the main objectives and findings)
- **Introduction** (setting out the purpose and aims of the survey, background to research, defines terms and concepts etc.)
- **Methodology** (describes method of sampling and information on survey population, data analysis and statistical procedures used)
- **Findings and analysis** (details of sample numbers, response rates, results and interpretation of tabulations)
- **Conclusions and recommendations** (summarising major findings and outline future actions)
- **Appendices and references** (i.e copy of the questionnaire).

11. Documentation

Documentation refers to;

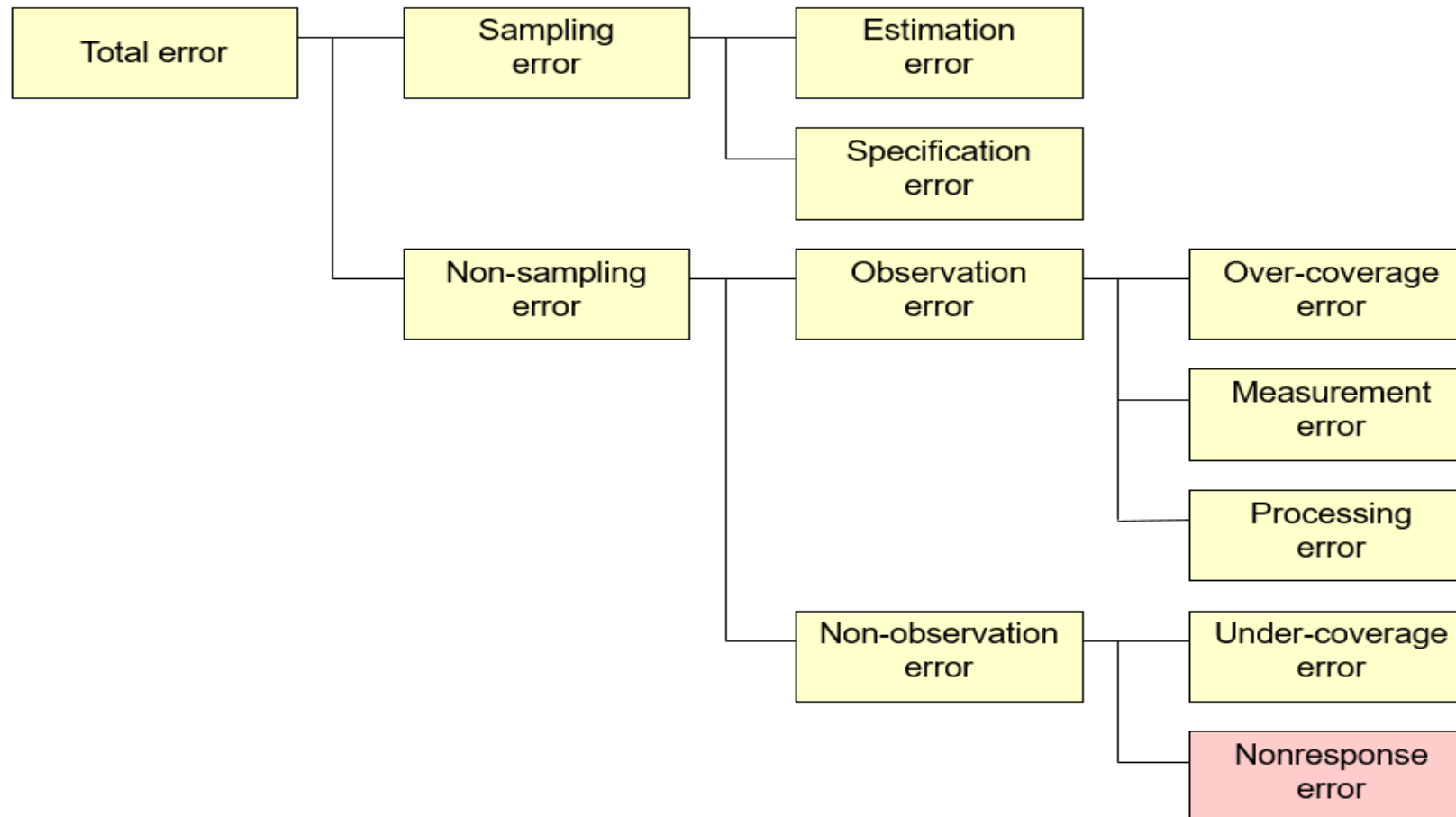
- ✓ the description of statistical activity including the concepts, definitions and methods used,
- ✓ the statistical production process itself with its information systems and work instructions.

By means of documentation users of statistics should be able to:

- Find out what statistics and statistical data (files) there are,
- Find the statistics and data they need,
- Interpret and analyse statistics,
- Process the data



Errors in surveys



Source: Cobben, F. (2009)

Sampling Errors

Sampling error arises from estimating a population characteristic by measuring only a portion of the population rather than the entire population.

- ✓ Estimation error: consequence of randomization.
- ✓ Specification error: sampling frame.

The magnitude of the sampling error can be controlled by,

- the sample size (it decreases as the sample size increases),
- the sample design
- the method of estimation.

The most commonly used measure to quantify sampling error is sampling variance.

If the variance is relatively large, then the estimate has poor precision and is unreliable.

Factors affecting the magnitude of the sampling variance include:

- The variability of the characteristic of interest in the population
- The size of the population
- The response rate
- The sample design and method of estimation

Nonsampling Errors

Nonsampling errors can be defined as errors arising during the course of virtually all survey activities, apart from sampling. They are present in both sample surveys and censuses.

- ✓ **Under-coverage:** mixed-mode data collection.
- ✓ **Measurement error:** questionnaire design, interviewer training, editing techniques, imputation techniques.
- ✓ **Processing error:** editing techniques, imputation techniques.
- ✓ **Nonresponse:** reduction techniques (contact strategy, interviewer training, refusal conversion), correction techniques (weighting, adjustment).



References

Franklin, S., & Walker, C. (2010). Survey methods and practices. Statistics Canada. *Social Survey Methods Division, Ottawa*. (Originally published in October 2003)

Cobben, F. (2009). Nonresponse in sample surveys: methods for analysis and adjustment. The Hague: Statistics Netherlands.

Introduction to Survey Design

- Defining the survey objectives, variables and concepts
- Defining the population to be surveyed
- Frames and Coverage

Cenker Burak METİN

Head of Survey and Questionnaire Design Group

Defining the survey objectives, variables and concepts

Poorly defined objectives will often lead to statistical activities that deliver data which do not meet the desired uses of the stakeholders and can lead to inefficiencies in the design of the activity.

For example, collecting information on income

- as a broad, cross classificatory variable expressed in a small number of broad ranges; or
- as a detailed dollar figure collecting only wage and salary income; or
- as detailed variables indicating dollar figures for all forms of income.

These different forms of income are collected to meet different intended uses of the data and have considerably different costs associated with their collection.



Balance the requirements of the user/users with the constraints of budget, resources and time available

Defining the survey objectives, variables and concepts

Suppose that a survey on poverty is to be conducted. It is not enough to indicate that the purpose of the survey is to provide information on, for example, ‘housing conditions of the poor’. Such a vague statement may serve as a broad description of the survey’s general theme, but ultimately it must be expanded into more specific language.

- What is meant by ‘housing conditions’?
- What precisely is meant by ‘poor’? Is poverty to be measured in terms of income, expenditures, debts, or all of these?



Concepts and Operational Definitions

For example, in a survey of businesses, respondents may need to be classified according to their industrial activity. If this is done using a standard format, the resulting statistics can be compared with existing published statistics or with a later repeat of the survey.

Likewise, in a household survey, if income is collected using the standard definition, it can be used on a comparable basis. By making use of standard concepts and data items it may also be possible to integrate data from different organisations. This integration improves the comparability and relevance of the data and reduces duplication of data collection.

- Concepts and variables should be clearly defined and linked to the research.
- Standard concepts, variables, classifications should be used.
- International definitions and classifications should be considered.
- Concepts, variables and classifications used in the research should be documented and differences from standards should be recorded.

Defining the population to be surveyed

✓ Target and Survey Populations

The target population or scope of the survey is the group about which inferences would like to be made from the survey data.

The following factors are essential in defining the target population;

- ❖ the type of units that comprise the population and the defining characteristics of those units (Who or What?)
- ❖ the geographical location of the units (Where?)
- ❖ the reference (time) period under consideration (When?)

The survey population is the population that is covered by the survey.

Frames and coverage

✓ Survey Frame

Once the client and statistical agency are satisfied with the definition of the target population, some means of accessing the units of the population is required. The survey frame (also called the sampling frame when applied to sample surveys) provides the means of identifying and contacting the units of the survey population.(Survey Methods and Practices, 2010)

The frame chosen determines the definition of the survey population and can affect the methods of data collection, sample selection and estimation, as well as the cost of the survey and the quality of its outputs.

✓ Selection of a Survey Frame

A statistical survey require the use of a correct sampling frame

The coverage, completeness, timeliness, information content and accuracy of the frame are critical factors

✓ Types of Frames

A list frame can be defined as a conceptual list or a physical list of all units in the survey population.

Examples of list frames are:

- Statistics register (e.g., a list of all births and or deaths in the population);
- Business register (e.g., a list of all businesses in operation);
- Address register (e.g., a list of households with civic addresses);
- Telephone directory (i.e., a list of all households with published telephone numbers);

An area frame is a special kind of list frame where the units on the frame are geographical areas. The survey population is located within these geographic areas.

Area frames may be used either when the survey is geographic in nature or when an adequate list frame is unavailable, in which case the area frame can be used as a vehicle for creating a list frame.

A multiple frame is a combination of two or more frames, (a combination of list and area frames or of two or more list frames).

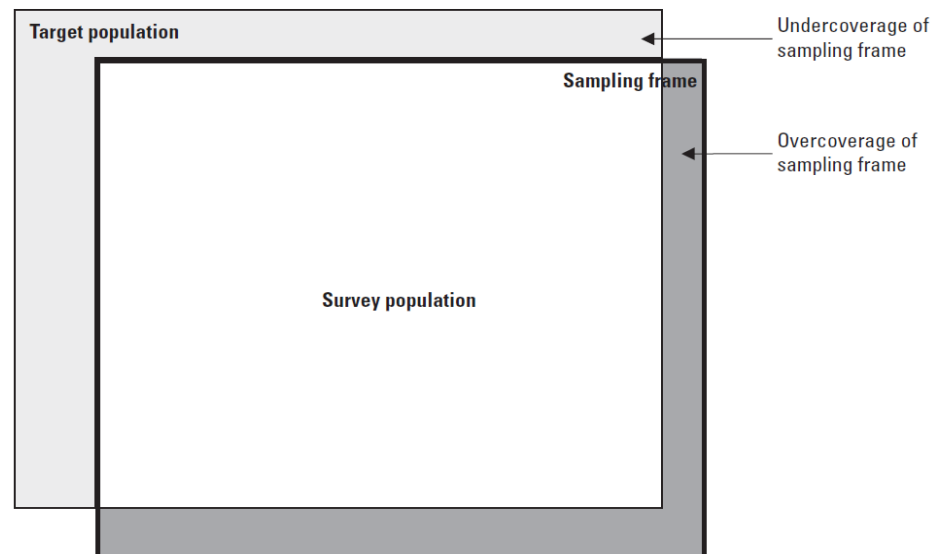
✓ Frame Defects

Undercoverage: exclusions from the frame of some units that are part of the target population.

Overcoverage: inclusions on the frame of some units that are not part of the target population

Duplication: the same unit appears on the frame more than once

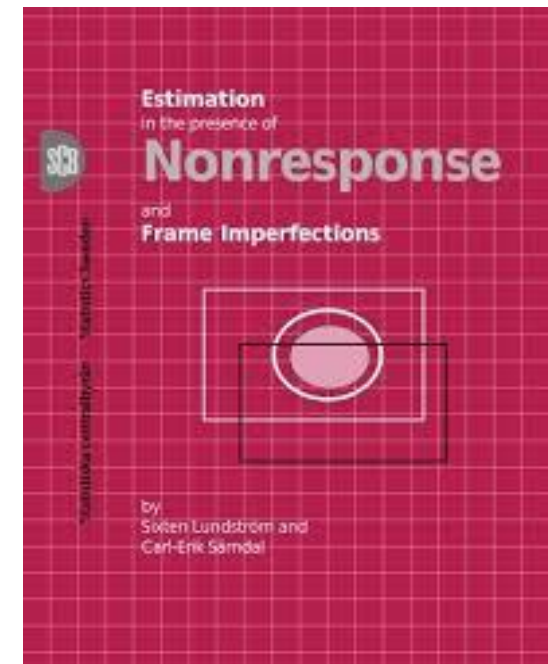
Misclarification: incorrect values for variables on the frame



Source: Quality Guidelines for Official Statistics (2002)

In order to choose and make the best use of the frame;

- ✓ assess different possible frames
- ✓ avoid using multiple frames
- ✓ use the same frame with the same target population
- ✓ incorporate procedures to eliminate duplication and to update frame
- ✓ emphasise the importance of coverage
- ✓ monitor the quality of the frame coverage periodically
- ✓ include descriptions in the survey documentation



References

Franklin, S., & Walker, C. (2010). Survey methods and practices. Statistics Canada. *Social Survey Methods Division, Ottawa*. (Originally published in October 2003)

Laiho, J., & Hietaniemi, L. (2002). Quality Guidelines for Official Statistics. Statistics Finland.

Introduction to Survey Design

- Data collection methods
- PAPI, CATI, CAPI and other methods
- Advantages and Disadvantages

Cenker Burak METİN

Head of Survey and Questionnaire Design Group

Data collection methods

Data collection is the process of gathering the required information for each selected unit in the survey.

- Expensive, time consuming, requires extensive resources
- Contributes to the image of the statistical agency
- Has a broad impact on the agency's relevance and on the quality of its data.

Regarding the method of data collection, many decisions must be made

- ✓ Should the questionnaire be administered by an interviewer?
- ✓ Should a combination of methods be used?
- ✓ Should the questionnaire be paper or computer-based?
- ✓ Should administrative data be used to collect some of the survey data?
- ✓ Should data collection for several surveys be combined?

Data collection methods

The method of data collection should be chosen to

- + Achieve high participation rate
- + Collect data as complete and accurate as possible
- + Minimize the response burden
- + Manage budget and operational constraints.

The other factors to consider are:

- Complexity of topic and nature of questions;
- Respondent preference;
- Sampling frame and target population;



The basic methods of data collection are:

- Self-Enumeration (PAPI, CASI)
- Interviewer-assisted (PAPI, CAPI, CATI)

Table 2.2 Taxonomy of Computer Assisted Interviewing methods¹ (de Leeuw and Nicholls, 1996)

Specific method	Computer assisted form
Face-to-face interview	Computer Assisted Personal Interviewing (CAPI)
Telephone interview	Computer Assisted Telephone Interviewing (CATI)
Self-administered form	Computer Assisted Self Interviewing (CASI), Computerized Self-administered Questionnaire (CSAQ)
Postal survey	Disk by Mail (DBM), Electronic Mail Survey (EMS)
Panel survey	Computer Assisted Panel Research (CAPAR), Teleinterview
Interviewer present	Computer Assisted Self Interviewing with Interviewer Present (CASI or CASIIP) Question Text on Screen: Visual (CASI-V) Text on Screen and on Audio (CASI-A)

¹ General name: CADAC (Computer Assisted Data Collection), CASIC (Computer Assisted Survey Information Collection), CAI (Computer Assisted Interviewing).

Source: Quality Guidelines for Official Statistics (2002)

PAPI, CATI, CAPI and other methods

Personal Interviews

The interview is implemented face-to-face with the respondents. (PAPI, CAPI)

Advantages of personal interviews :

- + Often have the highest response rates
- + High success rate to convert refusals
- + Interviewers may inspire confidence to respondents
- + Enable to make direct observations
- + Longer interviews than telephone

Disadvantages of personal interviews :

- Usually the most expensive method
- Difficult to find respondents at home or work
- Challenges for employing qualified interviewers
- Potential interviewer bias

PAPI, CATI, CAPI and other methods

Telephone Interviews

Interviewer assists the respondent by telephone. (PAPI, CATI)

Advantages of telephone interviews :

- + Are usually cheaper than personal interviews.
- + May be used for sensitive questions
- + Quality control of interviewing process may be implemented
- + Call-backs for people "not answering"



Disadvantages of telephone interviews :

- Restrictions for complexity and the number of questions
- Non-response may be higher than personal interviews
- Potential bias because of frame imperfections about phone numbers
- Confidentiality may be a problem

Computer-Assisted Data Collection

PAPI is still sometimes used, however it is more common to implement computer-assisted methods.

Advantages of **paper based** methods:

- + Are cheaper
- + Require less time to get ready

Disadvantages of **paper based** methods:

- Require one more step as data entry process
- The questionnaire should not have complicated skip patterns
- Confidentiality of the questionnaire forms

Computer-Assisted Data Collection

Computer-assisted methods have the main advantage of data collection and data entry simultaneously.

Advantages of **computer-assisted** methods:

- + Easier to monitor and control the quality of the data
- + Take advantage of automatic editing
- + Questionnaire may be more complex
- + Reduce response error for panel surveys

Disadvantages of methods:

- Require extensive and costly development work
- Basic knowledge of software application
- Confidentiality of respondent data
- Technical difficulties and problems



Other methods of data collection

- Direct observation
- Electronic Data Reporting
- Administrative Data
- Combining Methods



Question

What is the most used data collection method in your Statistical Office?

- | | |
|-----------------------------------|---|
| • PAPI (Face-to-Face) | 1 |
| • PAPI (Telephone) | 2 |
| • CAPI | 3 |
| • CATI | 4 |
| • Postal Survey | 5 |
| • Other (Please write the method) | 6 |

Operational process of data collection

Data collection process must be organized efficiently since it has the greatest cost in the survey..

Main components of the operational process and actions should be made

- ✓ Gain respondent co-operation (Good public relations get better reputation)
- ✓ Work with qualified interviewers
- ✓ Organization of the head and regional offices
- ✓ Preparation of data collection procedures
- ✓ Conducting interviews
- ✓ Monitoring quality and performance

References

Franklin, S., & Walker, C. (2010). Survey methods and practices. Statistics Canada. *Social Survey Methods Division, Ottawa*. (Originally published in October 2003)

Laiho, J., & Hietaniemi, L. (2002). Quality Guidelines for Official Statistics. Statistics Finland.

Questionnaire Design and Testing



Cenker Burak METİN

Head of Survey and Questionnaire Design Group

QUESTIONNAIRE DESIGN PROCESSES

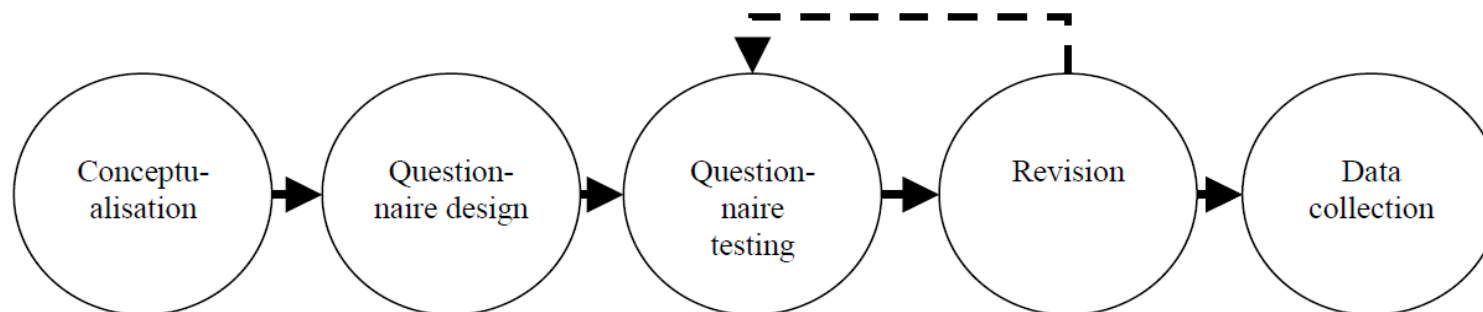
A questionnaire (or form) is a group or sequence of questions designed to obtain information on a subject from a respondent.

The questionnaire in the first instance is a measurement instrument. Its main purpose is to operationalise the user's information demand into a format which allows a statistical measurement.

A well-designed questionnaire should:

- ✓ collect data efficiently with a minimum number of errors and inconsistencies;
- ✓ be respondent friendly and interviewer friendly (if interviewer-assisted);
- ✓ lead to an overall reduction in the cost and time associated with data collection.

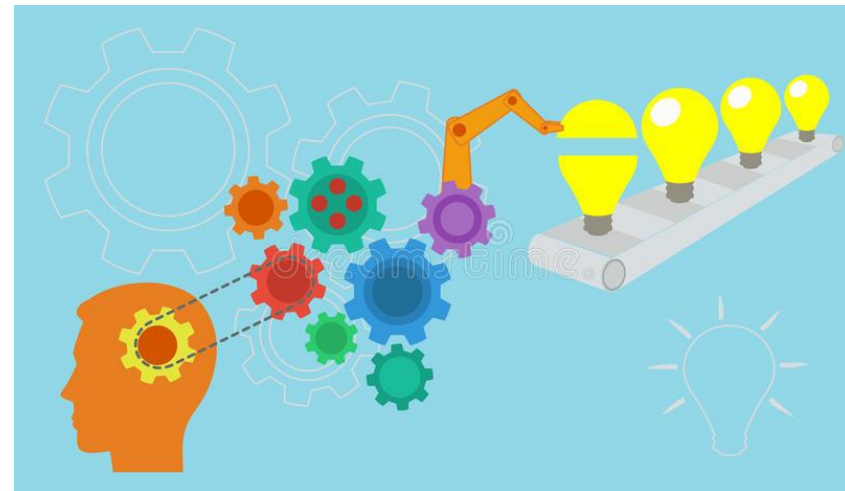
The five stages of questionnaire design and testing



1. Conceptualisation

Before one can even start to think about the wording of the questions, the conceptual basis of the questionnaire has to be specified and operationalised. Steps of conceptualisation;

1. Literature search
2. Survey objectives
3. Research strategies and methods for conceptualisation and operationalisation
4. Exploring concepts: focus groups and in-depth-interviews
5. Define variables and draft a tabulation plan
6. Data collection mode



2. Questionnaire Design

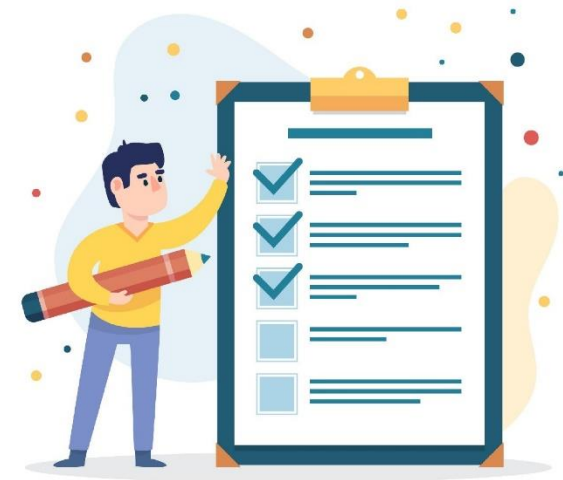
Principles of writing questions

- The respondent should:
 - ✓ clearly understand what he or she is being asked,
 - ✓ in principle be able to answer to the question, and
 - ✓ understand how the answer has to be given
- It is important to make sure that the actual respondents have the knowledge and necessary technical skill to answer the questions.
- They must understand what information they are expected to give and they must be able to find this information in their experience, memory, accounts, diaries, different activities, etc.

2. Questionnaire Design

Types of questions

- With regard to the information or data that can be obtained, there are four main types of survey questions:
 - ✓ Factual (Classification or Demographic Questions, Knowledge Questions)
 - ✓ Behavioural
 - ✓ Opinion (a person's attitude, sensitivity, validity of response)
 - ✓ Hypothetical questions.



2. Questionnaire Design

Questions formats

- Questions are classified as one of two types - **open or closed** - depending on the amount of freedom allowed in answering the question.
 - ✓ Open Questions (allow the respondents to answer the question in their own words)
 - Numeric Open- End
 - Text Open End
 - ✓ Closed Questions
 - Limited choice questions
 - Multiple choice questions
 - Partially closed questions provid



2. Questionnaire Design

Questions formats

- In choosing between these two alternatives (open or closed questions), consideration should be
 - ✓ given to factors such as the data requirement,
 - ✓ the kind of information required,
 - ✓ the level of accuracy needed,
 - ✓ processing facilities,
 - ✓ e.g. resources for doing the coding,
 - ✓ the position of the questions on the form
 - ✓ and the sensitivity of the question

2. Questionnaire Design

Response categories

- The most generally accepted principle about response categories states that it is very important to ensure that they are adequate, exhaustive and disjoint.
 - ✓ Number of responses options (no overlapping, long list should be avoided)
 - ✓ Order of response options
 - Can introduce bias.
 - Primacy and recency effects
 - “Don’t know” option should be included but put it to last
 - If some options are more socially desirable than others these should go last

2. Questionnaire Design

Question phrasing

- *Language* (appropriate language for what is asked, and provide the appropriate information)

Respondents

- ✓ have a limited vocabulary.
 - ✓ can understand positive instructions more easily than negative ones.
 - ✓ know nothing about National Statistics Institutes procedures or structure.
 - ✓ might not understand why the information asked for is needed.
- *Words* (simple and unambiguous terms, correct terminology, technical and negative words avoided)
 - .
 - *Sentences length*
 - ✓ Lengthy or complex questions can exceed the respondent's capacity
 - ✓ Short sentences are easier to understand than long ones.

2. Questionnaire Design

Question phrasing

- *Tone* (A change in wording can result in a change in responses)
 - ✓ “Do you think that gun ownership should be forbidden?” , or
 - ✓ “Do you think that gun ownership should not be allowed?”

Respondents

- ✓ Distort their responses to such statements, intentionally or unintentionally;
- ✓ Provide neutral responses (if available) or not respond to the statement question;
- ✓ Refuse to complete the entire questionnaire.
- *Order of clauses* (Sentences should have clauses in chronological order to aid comprehension)
- *Instructions*

2. Questionnaire Design

Defining the questionnaire flow

- For the overall readability of a questionnaire, a smooth progression through the questions is particularly important to minimise the nonresponse and measurement errors.
- The questions on a questionnaire should follow a sequence
- Two specific aspects related to sequencing
 - ✓ Order and grouping of questions
 - A preceding question can influence the attitude toward a following one.
 - It is recommended to keep similar topics together
 - The order of the sections is another important issue to consider.
 - The location of sensitive questions is also to be considered.
 - ✓ Filter Questions
 - A filter question is used to exclude respondents from subsequent questions if they do not apply
 - Filter questions are also used in interviewer based surveys
 - Electronic questionnaires can make complex sequencing much easier.

2. Questionnaire Design

Length

- The questionnaire length should be balanced considering the response burden, the mode of data collection and the fulfilment of survey goals.
- Towards the end of a long questionnaire, respondents may give less thought to their answers and concentrate less on the instructions and questions, thereby decreasing the accuracy of the information they provide (fatigue point effect).
- Whereas **a face-to-face interview should not exceed one hour**, the duration of **telephone interviews** and the completion of electronic questionnaires should only last **about 30 minutes**, and **45 minutes for mail surveys**. Of course, these limits can be exceeded when respondents are easy to motivate.

3. Questionnaire Testing

- Questionnaire testing is critical for identifying problems for both respondents and interviewers
- Questionnaire testing is a broad term that incorporates many different methods or combinations of methods.
- Two major categories of questionnaire testing methods – prefield and field methods.

✓ Pre-field methods

- **Expert groups** are composed of survey methodologists or questionnaire design experts, but also subject-matter experts
- **Observational interviews** are frequently used in order to identify problems in the wording, question order, visual design etc. of self-administered questionnaires. They also provide reliable estimates of the time needed to complete the questionnaire.
- **Cognitive interviews** are typically used after a questionnaire was constructed based on focus groups and has been improved in expert groups. The objective of cognitive interviews is to obtain qualitative information on how the questions are understood and answered by actual respondents.

3. Questionnaire Testing

- **Field methods** are those used to evaluate questionnaires tested under field conditions.
- Field testing often includes bigger sample sizes and allows quantitative analyses.
- The focus is more on the complete questionnaire instead of individual questions.
- Field methods include
 - **Behaviour coding** consists of systematic coding of the interaction between the interviewer and the respondent from live or taped interviews in order to evaluate the quality of the questionnaire.
 - **Interviewer debriefings** consist of an organised discussion of the questionnaire between interviewers who conducted the fieldwork, and the designers/researchers.
 - **Respondent debriefings** involve incorporating structured follow-up questions at the end of a field test interview or focus group style discussions with other interviewers to elicit quantitative and qualitative information about the respondents' interpretations of survey questions.

References

Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System, EUROSTAT, 2006

Questionnaire Development, Statistics Netherlands, 2012