

DATA COLLECTION



Learning Objectives

- ❑ To determine research design (*case study, Survey Multcase study*)
- ❑ To determine research techniques (*Quantitative or Qualitative*)
- ❑ To determine population (units of experiment) inquiring
- ❑ To identify Sampling techniques (*cluster, stratified*)
- ❑ To determine Sample size
- ❑ To identify data collection methods applied
- ❑ To determine sampling methods(s) / Techniques
- ❑ To identify types of data collected
- ❑ To determine the analytical methods used



Gather Information

- ❑ Conduct research to get the required data.
- ❑ Data can be collected both from primary and secondary sources.
- ❑ Primary sources of data are data from the original source; gathered orally (interview), in written form (questionnaire) or through observation or experiment or own experience.
- ❑ Secondary sources are written documents available in the library or through the internet.
- ❑ Apply the research techniques to get the information you need for the problem.

DATA COLLECTION METHODS



Questionnaire Design

The steps required to design and administer a questionnaire include:

1. Defining the objectives of the survey
2. Determining the sampling group
3. Writing the questionnaire
4. Administering the questionnaire
5. Interpretation of the results



Key Preparation

- ❑ Before you decide to design a questions, clearly articulate what problem or need is to be addressed using the information to be gathered by the questions.
- ❑ Review why you are doing the evaluation and what you ought to accomplish by it.
- ❑ The reason for evaluation determine the which information is relevant and therefore which questions to ask.



Directions to Respondents

- ❑ Include a brief explanation of the purpose of the questionnaire.
- ❑ Include clear explanation on how to complete the questionnaire.
- ❑ Note conditions of confidentiality.



Content of Questions

- ❑ Ask about what you need to know i.e. what you want to evaluate.
- ❑ Will the respondents be able to answer your questions? i.e. do they know the answer?
- ❑ Will the respondents want to answer the question? i.e. is it private or silly?
- ❑ Include both open and closed-ended questions.



Order of Questions Cont'd

- ❑ Include a question to get respondent's impressions of the questionnaire itself.
- ❑ Pilot or test your questionnaire on a small group to test it.
- ❑ Finalize the questionnaire.



Wording of Questions

- ❑ Use a familiar language
- ❑ Ask one question at a time, avoid the use of conjunction (and)
- ❑ Avoid the using ‘not’ in your questions if you are having respondents answer ‘yes’ or ‘no’
- ❑ If you use multiple choice questions, be sure your choices are mutually exclusive and encompass the total range of answers.



Wording of Questions Cont'd

- ❑ The questions should be clear and specific.
(often, sometimes, rarely)
- ❑ Avoid leading questions.
- ❑ Be careful about the use of adjectives e.g.
child-like/ childish.
- ❑ Avoid embarrassing questions



Order of Questions

- ❑ Be careful not to include so many questions that potential respondents are dissuaded from responding.
- ❑ Start with fact-based questions and then go on to opinion-based questions.
- ❑ Attempt to get respondents commentary in addition to their ratings.
- ❑ Do not ask hypothetical questions



Response Categories

- Consider the response categories available for closed-ended questionnaire type:
 1. Numerical ratings e.g. 1 2 3 4 5 (legend: 1= very low, 5= Very high)
 2. Fully anchored rating scales: (1/strongly agree, 2/agree, 3/neutral, 4/disagree, 5/strongly disagree).



Response Categories Cont'd

- Rankings (poor, good, very good, excellent)
- Checklist (check responses that are applicable)



Strength of Questionnaires

- ❑ Good for measuring attitudes and eliciting other content from informants.
- ❑ Inexpensive (mail & group administered questionnaires)
- ❑ Perceived anonymity by respondents may be high.
- ❑ High reliability and validity



Strength of Questionnaires

- Ease of data analysis for closed-ended items.
- Useful for exploration as well as confirmation.



Weaknesses of Questionnaires

- ❑ Must be kept short
- ❑ Reactive effects may occur i.e. informants only presenting socially desirable info.
- ❑ No response to selective items
- ❑ Response rate may be low for mail and email questionnaires



INTERVIEWS

- ❑ Interviews are used to pursue in-depth information around the topic.
- ❑ Interviews may be useful as a follow-up to certain responses to questionnaires.
- ❑ Before you start to design your interview questions and process, clearly determine the need/ problem to be addressed by the information gathered.



Preparation for Interview

- ❑ Choose a setting with little or no distraction.
- ❑ Explain the purpose of the interview
- ❑ Address terms of confidentiality.
- ❑ Explain the format of the interview.
- ❑ Indicate how long the interview will take



Preparation for Interview Cont'd

- Tell them how to get in touch with you later if they want to.
- Ask them if they have any questions before you get started.
- Don't count on your memory to recall their answers. Ask for permission to record the interview.(using a tape recorder or note taking)



Types of Interviews

1. Informal, conversational interview
2. General interview guide approach
3. Standardized, open-ended interview
4. Closed, fixed-response interview



Types of Interview Questions

1. Behavior: what a person has done/doing
2. Opinion/values: what a person thinks about the topic.
3. Feelings: Looking for feelings
4. Knowledge: to get facts about the topic
5. Sensory: about senses (seen, touched, heard, tasted or smelled)
6. Background/demographics: age, education,



Sequence of Questions

- Get the respondents involved in the interview as soon as possible.
- Ask about facts before feelings
- Ask questions about the present before past or future.
- End by allowing the interviewee to provide any other information that they prefer to add and their impressions of the interview.



Wording of Questions

- ❑ Wordings should be open-ended.
- ❑ Questions should be as neutral as possible.
- ❑ Questions should be asked one at a time.
- ❑ Questions should be worded clearly.
- ❑ Be careful asking ‘why’ questions.



Carrying out Interview

- ❑ Occasionally verify the tape recorder (if used) is working.
- ❑ Ask one question at a time.
- ❑ Attempt to remain as neutral as possible.
- ❑ Encourage responses
- ❑ Be careful about the appearance when note taking.



Carrying out Interview Cont'd

- Provide transition between major topics
- Don't lose control of the interview



Immediately After the Interview

- Verify if the tape recorder, if used, worked throughout the interview.
- Go through your notes, if taken, to fill in the gaps.
- Write any observations made throughout the interview. (place, time, nervousness, surprises etc)



Strengths of Interviews

- ❑ Good for measuring attitudes and most other content of interest.
- ❑ Allows probing and posing of follow-up questions by the interviewer.
- ❑ Can provide in-depth information.
- ❑ Closed-ended interviews provide exact information needed by the researcher.
- ❑ Useful for exploration and confirmation.



Weaknesses of Interviews

- ❑ In-person interviews are expensive and time consuming.
- ❑ Reactive effects
- ❑ Investigator effects may occur
- ❑ Interviewees may not recall important information.



Weaknesses of Interviews Cont'd

- ❑ Perceived anonymity by respondents may be low.
- ❑ Data analysis can be time consuming for open-ended questions
- ❑ Measures need validation.



Break



SAMPLING METHODS



Population, Element & Sample

- *Population*: is that group (usually people) about whom you want to be able to draw conclusions
- *Element*: The unit about which information is collected.
 - Typically an individual, but can
 - be family, organization, agency, school
- *Sample*: One or more elements (people) from a population

Parameter, Statistics, Representative Sample

- *Parameter*: Is a summary description of a given variable in a population e.g. Census-Age of population
- *Statistics*: Is a summary description of a given variable in a sample
- *Representative sample*: A sample which accurately reflects the distribution of relevant variables in the target population



Population Characteristics

- 50% sons, 50% daughters
- Average age of parents 67
- 20% live with parents, 80% live apart



Sampling Frame

- A list of all the elements from which the sample is drawn
- e.g. Families with a family member in jail due to crime





Probability vs. Non probability Sampling

- Probability – Sample is selected based on a random process (which means it is based on chance).

- Non probability– Sample is selected using some nonrandom process (not based on chance).



Types of Probability Samples

Enhances likelihood of getting a representative sample.

- ❑ Simple Random Sample
- ❑ Systematic Sample
- ❑ Stratified Random Sample
- ❑ Cluster Sample
- ❑ Haphazard sampling



Haphazard Sampling

The researcher has no control over who to sample e.g. where the population is massive (anybody/whole country) and the required respondents are few; or the cost of sampling in terms of time and money cannot be justified. For instance, a radio station that wishes to interview a few people on the street on some matter. They will simply send reporters to the street and ask them, say, to speak to 20 people among them, 7 adult men, 7 adult women, 3 girls and 3 boys. That would be a haphazard sample.

Simple Random Sample

Each element in the population (N=20) has an equal probability (chance) of being selected for the sample

Population		
Sue	Sally	Bonnie
Mary	Beth	Jennifer
Tom	Rebecca	
Richard	Scott	
Mike	Ben	
Kate	Sam	
Andrea	Bob	
Jill	Stan	40
Steve	Rill	

Systematic Sample

□ Every k th element in the list is chosen for inclusion in the sample

□ K is called the sampling interval

$$K = \text{population size} / \text{sample size}$$

population has 20 people. And you have resources to sample 5 people. $K = 20/5 = 4$

Sample every 4th person on the list. Start at random location



1.Sue	10.Sally	19.Bonnie
2.Mary	11.Beth	20. Jennifer
3.Tom	12.Rebecca	
4.Richard	13.Scott	
5.Mike	14.Ben	
6.Kate	15.Sam	
7.Andrea	16. Bob	
8.Jill	17.Stan	
9.Steve	18.Bill	

A.Abayo



Stratified Random Sample

- Strata – is a group of people who share a common characteristic.
- Examples of strata– race, gender, marital status.
- Population is first divided into two or more strata.
- From each stratum, a simple random sample is taken.
- The sub samples (samples from each stratum)

Sex: Female

Sue	Sally
Mary	Beth
Kate	Rebecca
Andrea	Bonnie
Jill	Jennifer



Sex: Male

Tom	Ben
Richard	Sam
Mike	Bob
Steve	Stan
Scott	Bill

Cluster Sampling

- Students
- Classes
- Schools
- School districts
- Cities
- Counties



Accidental/Convenience/

Availability Sample

~~Convenience samples:~~

- This is sometimes seen as quite similar to purposive and haphazard samples. Perhaps the only difference is that the selection is based on practical convenience.
- The researcher is therefore more thoughtful than in haphazard samples. For instance, in a study of school children in the country, it may be convenient (for logistical, financial reasons, etc) to select the sample from schools in the researchers home city alone.
- Select people because they are readily available.
- Sample members chosen because convenient. Easy to access



Examples of Accidental Samples

- Place ad in newspaper. Those who call to volunteer to participate.

- People who belong to an organization



Quota Sampling

- Divide the population into various categories
- Determine the number of people to be selected for each category
- For each category, you obtain an accidental sample until you fill the quota of people you need for that category



Purposive or Judgmental Sample

- Select your sample on the basis of your knowledge of the population and nature of your research aims.

□ .



Purposive/Judgmental Sample

E.g. a researcher on education quality in schools may believe that the data they seek is best obtained from **head teachers of schools in urban areas** hence choose the sample from this group rather than all head teachers.

the sample is purposely selected for Possessing certain characteristics



Snowball Sample

- Find someone who fits the criteria for the study (accidental/convenience sample)

- Interview person and at end of interview, you ask if he/she knows of other people who meet the study criteria and may be willing to participate

Present the Findings and Analysis

Usually it is expected that all the considerations will be on:

- Presentations of the findings
- Data analysis
- Interpretations of findings *(in relation to objective(s) and purpose of the study)*

DATA QUALITY

Quality of Data - Reliability and Validity

- These two terms are often used to describe the quality of methods for data collection. What do they mean? [\[1\]](#)
- Reliability is the extent to which a method/tool will produce the same result if applied two or more times.
- Imagine a thermometer that reads 90 degrees C when it should read 100 degrees C in boiling water at sea level. It would be considered 'unreliable'. The same is true when the thermometer is applied three times and it reads 100, 85 and 110 degrees C.
- [\[1\]](#) Here, we will not go into a detailed scientific explanation of the terms.

Checking Reliability

There are especially two ways to check the reliability:

- 1) use the same method/tool at two points in time and compare the consistency of the results and
- 2) Several observers collect the data at the same time. This tells you whether two people are seeing and interpreting the same thing in the same way.



Validity

- Validity is the degree to which a method/tool measures what you want it to measure.

Quality of Data

- ❑ Monitoring data are never completely clear-cut and absolute.
- ❑ Data collection is always flawed in some way, and there are always questions of reliability and validity.
- ❑ Error-free instruments do not and cannot exist in the measurement of complex human, social, economic, political and behavioural phenomena.
- ❑ ... Monitoring is only of use if one believes that some systematic information is better than none.
- ❑ ... Monitoring does not provide final answers, but it can provide directions. (Adapted from Michael Quinn Patton)

Relation between reliability and validity

- Validity is more important than reliability.
- If a method/tool is unreliable it cannot be valid. Reliability is a necessary but not a sufficient condition for validity.
- A method/tool may produce the same results consistently; however, it may not measure what you intend to measure.

Example: A scale may tell you consistently the weight of a Mango, but a) the scale may be – consistently - wrong and b) weight is not a valid indicator for e.g. quality or sweetness of the fruit.

Reliability and validity

- **Reliability:** This is the degree to which data collected are consistent among different observers or the same observer at different times.
- **Validity:** The extent to which the information which has been collected can yield the desired reality as it is in a given situation.



Representativeness

- **Representativeness** refers to the question of whether the group or situation being studied are typical of others.
- If a group or situation is representative of others, then researchers can **generalize** their findings.
- Representativeness is determined by representative sources of data.



Objectivity

- ❑ As a researcher you should not be biased or prejudiced in the way you carry out investigation.
- ❑ As a researcher, you should avoid letting your values, beliefs and pre-existing ideas affect or influence the way in which you develop a project, conduct research or analyze the data.
- ❑ You need to justify the conclusions that you arrive at and evaluate whether the conclusions that you draw are valid and reliable.



Quality control

- There are different ways of ensuring quality control in data collection. These include:
 1. Triangulation of methods and multi-disciplinary teams.
 2. Thorough training of the field teams.
 3. Checking through the questionnaires with the interviewers at the end of each day's fieldwork.
 4. Pre-testing of the tools for data collection before fieldwork.



Triangulation and Quality Control

- Triangulation refers to the use of at least three diverse but appropriate and complementary methods or techniques to either collect information from a given situation or analyze it.

- Methods, sources of information and multi-disciplinary teams can be triangulated in data collection.

Triangulation

Triangulation and quality control help in ensuring the following:

1. Completeness of data
2. Balance and objectivity in an evaluation
3. Reliability of the information and its subsequent analysis.
4. Validity of the results

Quality control

The Fieldwork Team

Recruitment

- ❑ The field team should be recruited considering the following criteria.
- ❑ Openness and friendliness
- ❑ Politeness and sensitivity to issues
- ❑ Personal appearance: dress, etc.
- ❑ Basic education (though not always) is needed
- ❑ If possible they should be local people from the area where the evaluation will be carried out.



Quality control

The field work team

Training

- ❑ Every fieldwork team member needs training on the tools to be used. At least 2-3 days are needed for training.
- ❑ Training will also give an opportunity to test field recruits on their capability to collect data.
- ❑ If possible a training manual should be prepared for the field team.
- ❑ During the training of fieldwork team, the following tasks should be accomplished.
- ❑ Introduce the background of the evaluation to them
- ❑ Give information about the agency/program to be evaluated.



Quality control

Training of the field work team

- ❑ Give information about the agency/program to be evaluated.
- ❑ Give the rationale for the evaluation.
- ❑ Give a general brief on the terms of reference (TOR).
- ❑ Explain the methodology to be used during the evaluation.
- ❑ Emphasize the need for quality control.
- ❑ Explain the importance of the information to be collected
- ❑ Discuss how they will introduce the evaluation to the target audience.
- ❑ If possible translate the questions into the local language.
- ❑ Conduct mock interviews to assess the validity of the questions.

Quality control

Pre-testing the Questionnaire

- The purpose of pre-testing is to identify **weaknesses, ambiguities** and **omissions** before finalizing your questionnaire. It improves the quality of the questionnaire.
- Pre-testing will help to find out whether the questions are well presented, clearly understood and easy to answer and do not cause anxiety, embarrassment, resistance, etc.
- Pre-testing also shows how long it will take to interview all the respondents and to indicate the best times for interviewing.

Quality control

Pre-testing the questionnaire Casley advances the questions below in relation to pre-testing.

- *Wording of questions:* Was the wording of the questions clear to the respondents? Did all respondents derive the same meaning from the questions?
- *Construction of sentence:* Were the sentences appropriate? Were they too short or too long? Did they give unnecessary details which confused the respondents?
- *Question format:* Was the question format suitable? If there were open-ended questions, was there a great variation in responses which would make them difficult to code? Were the existing response categories for the closed questions adequate in

Pre testing the questionnaire

Pre testing the questionnaire

- *Difficult questions:* Were there questions which the respondents found it difficult to answer?
- *Same answer:* Were there questions for which all the respondents gave the same answers?
- *Refusal rate:* Was there a tendency for respondents to refuse to answer particular questions?
- *Time requirement:* What was the approximate time needed to complete the questionnaire? Did the respondent seem to tire at the end?
- *Interviewer's Convenience:* Did the enumerators find it difficult to administer some parts of the questionnaire? Were additional instructions needed? Were the probes appropriate?
- *Coding:* Were there problems in coding the data?
- *Usefulness of the data:* Was the questionnaire able to

Bias

-
- Bias is a systematic error that decreases the quality of information.
 - What factors might introduce bias into data collection?

Potential sources of Bias in data collection

- Awareness: People aware that they are part of a monitoring activity may respond in uncharacteristic ways. Some are anxious others become defensive.
- Role selection: The awareness of being measured may influence people to play a special role. They ask, “What is expected of me?” and act accordingly.



Potential sources of Bias in data collection

- Yeah-factor: People are much more likely to say “yes” than “no” to any particular question.
- Social desirability: People tend to answer a question the way they think you - as the data collector - wants the question answered, rather than according to their true feelings or opinions.

Potential sources of Bias in data collection

- Population stability over areas: The same way of collecting data in two different geographic areas may tap different kinds of people. They may differ in demographic, ethnic or other characteristics.
- Content stability over area: Especially large integrated programmes are not uniform in content and activities throughout the area where they are implemented. Therefore, monitoring has to take this diversity into consideration.

Potential sources of Bias in data collection

- Professional and personal bias: We've often undergone professional training and acquired specialized expertise. In addition, we have our own worldview about how things are interrelated. Therefore, we are prone to see the world from our - restrictive - perspective.
- 'Tarmac' bias: Areas chosen and visited for data collection tend to be those that are more accessible – e.g. near roads and settlements.

Potential sources of Bias in data collection

- Season and time bias: Data collection has its own seasonality and is often confined to the dry season after the rains. This may lead to under- or over reporting of certain seasonal incidents (like sickness etc.). In addition, if we visit a village, farm or trading centre always during certain daytime, we may underrepresented groups who are not around for various reasons.
- Any data collection activity is subject to some forms of systematic error. But, it is impossible to control all sources of bias. But you can try to minimise each source by carefully planning the data collection (when, who, where, which tool(s) etc.)

QUALITATIVE DATA ANALYSIS



What is qualitative data

- Qualitative data is information gathered in a nonnumeric form.
- Common examples of such data are:
 - Interview transcript
 - Field notes (notes taken in the field being studied)
 - Video
 - Audio recordings
 - Images
 - Documents (reports, meeting minutes, e-mails)
 - Text
 - Written words
 - Phrases
 - Symbols describing or presenting people, actions and events in social life



What is Qualitative Data

- For example, Someone's interpretation of the world,
 - Why they have that point of view,
 - How they came to that view,
 - What they have been doing,
 - How they conveyed their view of their situation,
 - How they identify or classify themselves and others in what they say,

What is Qualitative Data Analysis?

- Qualitative Data Analysis (QDA) is the range of processes and procedures whereby we move from the qualitative data that have been collected into some form of explanation, understanding or interpretation of the people and situations we are investigating.
- QDA is usually based on an interpretative philosophy.
- The idea is to examine the meaningful and symbolic content of qualitative data.



Qualitative data analysis

- ❑ Except from occasional content analysis, qualitative data is rarely analysed using statistical analysis.
- ❑ This does not mean that the analysis is based on speculation or vague impressions.
- ❑ It is systematic and logically rigorous although in a different way.
- ❑ It is a systematic and a step-by-step approach.



What is Qualitative data analysis

- ❑ Analysing qualitative data is not a simple or quick task.
- ❑ Done properly, it is systematic and rigorous, and therefore labour-intensive and time-consuming.
- ❑ Good qualitative analysis is able to document its claim to reflect some of the truth of a phenomenon by reference to systematically gathered data," in contrast,
- ❑ Poor qualitative analysis is anecdotal, unreflective, descriptive without being focused on a coherent line of inquiry.
- ❑ At its heart, good qualitative analysis relies on the skill, vision and integrity of the researcher doing that analysis,



What is QDA

- ❑ Less emphasis on measurement, standardization and mathematical techniques.
- ❑ Focus on saturation as opposed to Quantitative data analysis which focuses on statistical guidance.
- ❑ Can be done in the field during data collection in the field.
- ❑ Can be done manually or using computer



How useful is computer analysis

- A computer package may be a useful aid when gathering, organising, and reorganising data and helping to find exceptions, but no package is capable of perceiving a link between theory and data or defining an appropriate structure for the analysis.
- To take the analysis beyond the most basic descriptive and counting exercise requires analytical skills in moving towards hypotheses or propositions about the data.

Data analysis during data collection

- ❑ In qualitative analysis the analytical process begins during data collection as the data already gathered are analysed and shape the ongoing data collection.
- ❑ This sequential analysis or interim analysis has the advantage of allowing the data collector and analyst to go back and refine questions, develop hypotheses, and pursue emerging avenues of inquiry in further depth.
- ❑ Crucially, it also enables the researcher to look for deviant or negative cases

Quantitative and Qualitative

Analysis: Similarities and Differences

Similarities

- ❑ Inference
- ❑ Involves a systematic Method/Process
- ❑ Comparison of gathered evidence to existing evidence
- ❑ Avoidance of errors, false conclusions and misleading inferences



Quantitative and Qualitative Analysis: Similarities and Differences Cont'd

Differences

Quantitative investigators

- Choose from specialized/standardized set of data analysis techniques
- Collect/condense all data into numbers before analysis commences

Qualitative Investigators

- Data analysis is less standardized

Differences Between Quantitative and Qualitative Data Analysis

Quantitative Investigators

- Manipulate numbers to test an abstract hypothesis with variable constructs

Qualitative Investigators

- Begin analysis while still collecting data
- Look for patterns/relationships
- Results of early analysis guide subsequent data collection



The process

- Data reduction-manipulating as per objectives, integrating data from different methods and making it orderly as per its importance. All data must be read.
- Data organization using themes/points/categorization
- Data interpretation-drawing conclusions, patterns, answering evaluation questions, reaching saturation.

Concept Formation in qualitative data analysis

- ❑ Qualitative analysis uses nonvariable concepts or simple nominal-level variables.
- ❑ These are in the form of general ideas, themes or concepts.
- ❑ Concept formation is an integral part of data analysis and begins during data collection.
- ❑ Conceptualization is one way of organizing and making sense of qualitative data.



Conceptualization

- ❑ Data is analyzed by organizing it into categories on the basis of themes, concepts or similar features.
- ❑ The analyst develops new concepts, formulates conceptual definitions, and examines the relationships among concepts.
- ❑ Eventually ,concepts are linked either as oppositional sets or as similar categories which are interweaved into theoretical statements.



Conceptualization

- ❑ Concepts are formed as you read through field notes and ask critical questions.
- ❑ Data is organized and ideas applied simultaneously to create or specify a case.
- ❑ Making or creating a case brings the data and theory together.

Coding into themes

- ❑ Looking for themes involves coding.
- ❑ This is the identification of passages of text (or other meaningful phenomena, such as parts of images) and applying labels to them that indicate they are examples of some thematic idea.
- ❑ At its simplest, this labelling or coding process enables researchers quickly to retrieve and collect together all the text and other data that they have associated with some thematic idea so that they can be examined together and different cases can be compared in that respect.



Coding

- Most typically, when coding, the analyst has some codes already in mind and are also looking for other ideas that seem to arise out of the data.
- When coding ask the following questions about the data you are coding:
 - "What is going on?"
 - What are people doing?
 - What is the person saying?
 - What do these actions and statements take for granted?
 - How do structure and context serve to support, maintain, impede or change these actions and statements?"



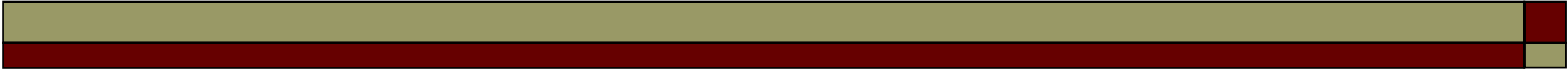
Ordering and coding of data

We will discuss two types of qualitative data:

- answers to open questions, and
- more elaborate narratives from loosely structured interviews or FGDs.

Answers to open questions

- (1) A first, basic step in the analysis of answers to open questions is to **list** the answers of a *sample of 20-25 informants* as they were provided
- (2) Then **read** the answers carefully, remembering the purpose of the question. For example The question ‘why are you smoking’ was supposed to help nursing students to develop an intervention against smoking.
- (3) **Make rough categories** of answers that seem to belong together and **code** them with a key word. For example, answer (It gives me pleasure) and answer (I like to blow smoke rings) could be labelled with the term ‘pleasure’, which could be abbreviated with the code *pleas*.
- (4) Then **list** again all answers but now **per code**, so that you get some 5-7 short lists, for example:

- 
-
- 5) Then **interpret** each list, and end up with some 5-7 meaningful categories with a characteristic key word. For example:
Pleasure, being sociable, giving status, giving self-confidence, addiction, defiance. There may be discussion on the need to split up some categories or combine others with few answers.



(2) Elaborate narratives

- ❑ The data from interviews with key informants or focus group discussions (FGDs) are as a rule more bulky than answers to open questions.
- ❑ The carefully transcribed field notes and tapes may consist of pages of narrative text.
- ❑ When analysing the texts we usually discover that, no matter how good our guidelines for the discussion were, the data contain valuable information but also a number of less essential details.
- ❑ In addition, the data is usually not presented in the order we need for our analysis, since informants may jump from one topic to the other.



Elaborate narratives

- To make the analysis easier, we have to **order** and **reduce** the data.
- Ordering is best done in relation to the objectives and the discussion topics.
- (1) **Reread** your objectives and discussion topics
- (2) Carefully **read** a number of the interviews, FGDs or narrative observations you want to process.

Example

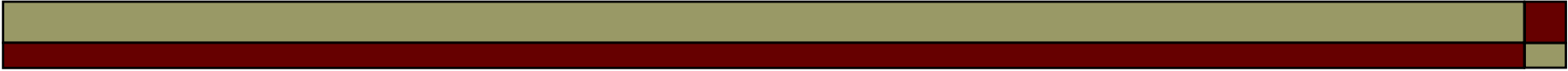
- In a gender and leprosy study it appeared that the discussion topic *stigma* had to be differentiated according to different social settings in which it occurred: among close relatives (parents-children), spouses, in-laws, and community members.
- Further, a distinction had to be made between self-stigmatisation (e.g., a wife diagnosed as a leprosy patient encouraging her husband to marry a second wife in order to prevent divorce, or a patient not attending community meetings for fear of being avoided) and stigmatisation by others.
- Different degrees of severity in stigmatisation could also be distinguished, varying from slight avoidance to complete expulsion.
- If stigma would be topic ,in your discussion list, you would mark everything related to stigma and add *key words* such as *self-stigm.*, *spouse*, *in-laws*, *comm.*, in the margin, as well as key words such as *sleep(ing) sep(arately)* or *divorce* indicating the severity of the stigma.

-
- (3) **List all key words that belong to a certain topic in the sub-categories** that have been developed under (2). E.g., everything belonging to stigma could be subdivided and listed in the four major social settings in which stigma was found to manifest itself.
 - (4) **Interpret** the data, e.g., distinguish the major forms in which stigma manifests itself in these different social settings, try to make a ranking order of severity and link it to other variables (such as degree of deformity, socio-economic status) in order to understand differences in stigma.
 - (5) Then **code all your qualitative data** in this way. If necessary, adapt your coding scheme as you order, code and interpret more data. In that case, you should again read and possibly re-code the material you have already processed.
 - **The information can be summarised into coding sheets COMPILATION SHEETS.**



Summarising data in compilation sheets

- summarising all data of each study unit per study population on separate compilation sheets.

- 
-
- Each interview, FGD or observation gets a number and is successively entered in that sequence on the relevant compilation sheet.
 - If there are different categories of informants within one study population, for example, young mothers and an older generation of mothers, or male and female patients, the data for these groups are entered on separate sheets.
 - If the topics covered in those sub-groups are not completely identical, it is important to be systematic and follow roughly the same sequence of topics for each category of informants.
 - The information inserted is summarised in key words and key sentences, clear enough to remember the statements informants made.
 - (As the number of each study unit is entered in the compilation sheet, it is always possible to go back to the original data and present the full statement, for example in a presentation or in the research report).

Further summarising of data in matrices, figures and tables

- **Matrices**
- Matrices can be used for quantitative as well as qualitative data comparison.
- In qualitative data we may compare different groups or data sets on important variables, presented in key words.
- A **MATRIX** is a chart that looks like a cross-table, but contains words (as well as, sometimes, numbers).



matrices

- ❑ Matrices facilitate data analysis considerably.
- ❑ They are the most common form of graphic display of qualitative data.
- ❑ They can be used to order and compare information in many ways, for example, according to:
 - ❑ time sequence (of procedures being investigated in different periods, for example),
 - ❑ type of informants (as in the example above), or
 - ❑ location of data collection (to visualise differences between rural and urban populations).



Diagrams

- Diagrams, like matrices, can be of great assistance in providing an overview of the data collected and in guiding data analysis.
- A DIAGRAM is a figure with boxes containing variables and arrows indicating the relationships between these variables.

Flow charts

- FLOW CHARTS are special types of diagrams that express the logical sequence of actions or decisions.
- Flow charts are especially useful to summarise different flows of events that are mutually connected.



Tables

- **Tables**

- A TABLE is a chart with rows and columns that has numbers in the various cells or boxes.

DRAWING AND VERIFYING

CONCLUSIONS

- Drawing and verifying conclusions is the essence of data analysis. It is not an isolated activity, however.
- When we start summarising our data in compilation sheets, flowcharts, matrices or diagrams, we continuously draw conclusions, and modify or reject quite a number of them as we proceed.
- Writing helps generate new ideas as well.
- Therefore **writing should start as early as possible**, right from the onset of data processing and analysis, if only for ourselves.
- No creative insights should get lost!

DRAWING AND VERIFYING

CONCLUSIONS

- **Note:**
- Collection, processing, analysis and reporting of qualitative data are closely intertwined, and not (as is the case with quantitative data) distinct successive steps.
- It may often be necessary to go back to the original field notes and verify conclusions, collect additional data if available data appear controversial, and get feedback from all parties concerned.



REPORTING QUALITATIVE DATA

- One way is summarising the major qualitative results in a separate section of the findings, with examples and quotations, following the objectives that guided the collection of this particular data.
- The results would then be discussed in the chapter ‘Discussion’, together with the results of other, more quantitative data collection tools and would subsequently be reflected in the summary of the findings and the recommendations.



REPORTING QUALITATIVE

- Another possibility is to fully integrate different data sets in the chapter of findings, ordered according to the objectives of the entire study.
- If quantitative and qualitative data have been analysed and sometimes even collected in an integrated way, it would also be logical to present them in an integrated fashion.
- Attention should be paid that no valuable data get lost.



STRATEGIES FOR TESTING OR CONFIRMING QUALITATIVE FINDINGS TO PROVE VALIDITY

1. Check for representativeness of data.

- Check whether you have indeed interviewed all categories of informants needed to get a complete picture of your topic (not relying excessively on talkative authorities).
- Make sure that you do not generalise from unrepresentative events.



STRATEGIES FOR TESTING OR CONFIRMING QUALITATIVE FINDINGS TO PROVE VALIDITY

2. **Check for bias** due to observer bias or the influence of the researcher on the research situation.

STRATEGIES FOR TESTING OR CONFIRMING QUALITATIVE FINDINGS TO PROVE VALIDITY

3. Cross-check data with evidence from other, independent sources.

- Actively cross-checking data, looking for independent evidence or counter-evidence, is one of the most important ways to enhance the validity of research data.
- **For example**, answers of husbands and wives (and other informants concerned) should confirm each other on such issues as who decides whether and what family planning methods should be used, who decides whether daughters should be circumcised, or what has changed in husband-wife relationships after the diagnosis of leprosy or another feared disease in one of the spouses.

STRATEGIES FOR TESTING OR CONFIRMING QUALITATIVE FINDINGS TO PROVE VALIDITY

4. Compare and contrast data

- If we want to be sure, **for example**, that variable A (high level of education) influences variable B (use of family planning methods) we have to compare a group of mothers with high education to a group of mothers with low education on their use of family planning methods.
- Comparing and contrasting data is important if you are attempting to **identify** your variables as well as to **confirm** associations among variables.

STRATEGIES FOR TESTING OR CONFIRMING QUALITATIVE FINDINGS TO PROVE VALIDITY

5. Use extreme (groups of) informants to the maximum.

- During study design and sampling it may be useful to look for categories of informants that represent the extremes on a certain variable.
- **For example**, you may find it most useful to study ‘drop-outs’ and regular attendees of TB services, leaving out the category of irregular attendees. This may be the most efficient way of identifying the key variables that influence the attendance behaviour of TB patients.

STRATEGIES FOR TESTING OR CONFIRMING QUALITATIVE FINDINGS TO PROVE VALIDITY

6. Do additional research to test the findings of your study.

- The results of your study may be so intriguing that you decide to do a follow-up study afterwards.
 - Such a study may be undertaken for several reasons:
 - to replicate certain findings,
 - to rule out (or identify) possible intervening variables,
 - to rule out rival explanations by investigating them, or
 - to look for negative evidence.
- Additional studies undertaken for one or more of these reasons may serve to make the results of your original study more convincing.



STRATEGIES FOR TESTING OR CONFIRMING QUALITATIVE FINDINGS TO PROVE VALIDITY

7. Get feedback from your informants.

- This is important not only for ethical reasons or because it will improve the chances that the results will be implemented, but also because **it will improve the quality of your study design, of your data, and of the conclusions drawn from these data.**
- Suggestions and additional information collected during feedback sessions will invariably increase the quality of your research report.

Tips to improve your analysis of qualitative data:

- Carefully review all the data.

- Organize comments into similar categories, e.g., concerns, suggestions, strengths, weaknesses, similar experiences, program inputs, recommendations, outputs, outcome indicators, etc.
- Try to identify patterns, or associations and causal relationships in the themes, e.g., all people who attended programs in the evening had similar concerns, most people came from the same geographic area, most people were in the same salary range, processes or events respondents experience during the program, etc.
- Try to combine the results of the quantitative and qualitative data.
- It is important to keep all documents for several years after completion in case they are needed for future reference.

