

Data Collection and Data Analysis

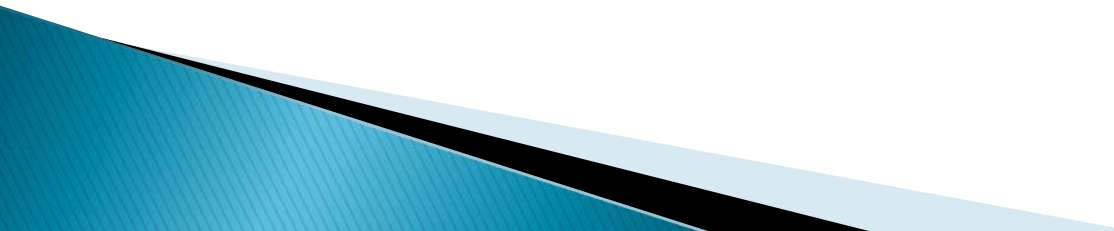
Community Empowered Research Training (CERT)

Yumary Ruiz, PhD, MPH

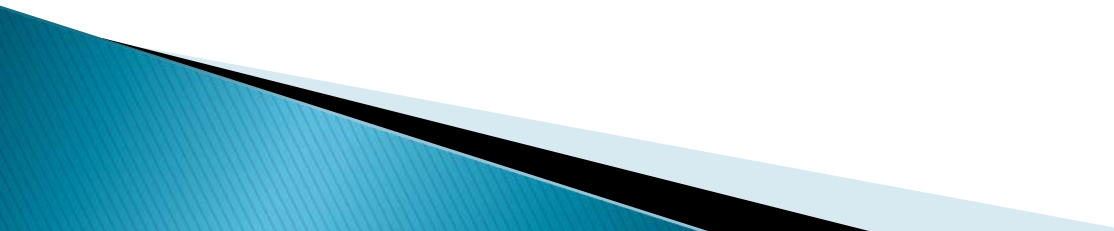
May 20, 2011



LEARNING OBJECTIVES

- ▶ Explain the difference between primary and secondary data
 - ▶ Describe the variety of methods used to collect both qualitative and quantitative data
 - ▶ Understand why sampling is necessary
 - ▶ Recognize various sampling methods
 - ▶ Become familiar with the basics of compiling a questionnaire
 - ▶ Discuss how quantitative and qualitative data are analyzed
- 

Data

- ▶ Are collections of pieces of information that can be organized in a way to increase knowledge
 - ▶ Types of Data: Primary, Secondary
 - **Primary data** contains new information that has not been previously analyzed
 - **Secondary data** are data that have already been collected by someone else
- 

Selecting a data collection method

- ▶ Factors that influence selection of a data collection method include:
 - availability of project and population resources
 - amount of time participants are willing to give to provide requested information
 - ability to recruit the target population
 - population's skills
 - training and skills of the staff
 - time available for data collection
 - space needs
 - questions
 - types of questions to be answered
 - depth and detail needed

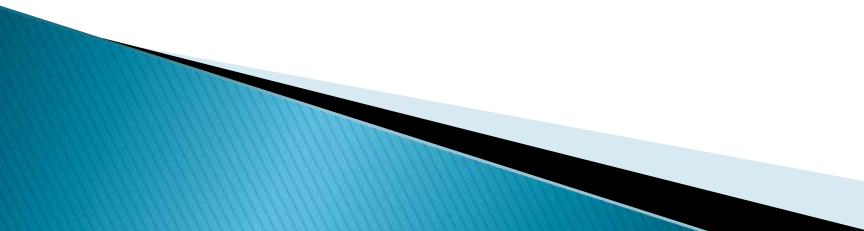
Qualitative vs Quantitative Data

- ▶ Quantitative
 - Researcher's judgement essential before and after
 - Creates variables that are amenable to statistical analysis
 - Logic rests on generalizability & representativeness
 - Sample size is criterion for judging rigour
- ▶ Qualitative approaches
 - Researcher "becomes the instrument"
 - Designed to best reflect experiences
 - Uses discourse as data, long descriptive text using people's own words and dialogues.
 - Logic rests on notice of saturation – the point at which no new insights are likely to be obtained
 - Saturation guides sample size

Quantitative Data Collection Methods



Collecting Quantitative Data

- ▶ Statistical data (secondary analysis)
 - ▶ Survey research (all require a structured questionnaire, where each observation unit is posed with the exact same questions)
 - mail questionnaire survey
 - e-mail questionnaire survey
 - face-to-face questionnaire survey
 - telephone questionnaire survey
- 

SECONDARY DATA SOURCES: *National*

- ▶ **American FactFinder** [<http://factfinder.census.gov/>] This website is the Census Bureau's primary vehicle for distributing the 2000 Census of Population and Housing, 1997 Economic Census, and American Community Survey. Currently includes 1990 Census Quick and Detailed Tables, American Community Survey
- ▶ **Center for Disease Control Data and Statistics** [<http://www.cdc.gov/DataStatistics/>]
- ▶ **ChildStats.gov** [<http://www.childstats.gov/americaschildren/>] Summary information on many issues surrounding children and children's health in the United States.
- ▶ **Data Resource Center for Child & Adolescent Health** [<http://www.childhealthdata.org/content/Default.aspx>] Access site for public data on the health and health-related services for children, youth, and families in the United States.
- ▶ **Department of Homeland Security Immigration Statistics** [<http://www.dhs.gov/ximgtn/statistics>] Data on refugees, permanent residents, and naturalized citizens.
- ▶ **Environmental Protection Agency: Office of Children's Health Protection & Environmental Education** [<http://yosemite.epa.gov/ochp/ochpweb.nsf/content/homepage.htm>]
- ▶ **National Health and Nutrition Examination Survey** [<http://www.cdc.gov/nchs/nhanes.htm>]
- ▶ **National Center for Health Statistics Data Warehouse** [<http://www.cdc.gov/nchs/datawh/ftpserv/ftpdata/ftpdata.htm>] Provides free public access to data sets, documentation, and questionnaires from NCHS surveys and data collection systems. Downloading instructions are available in "readme" files.
- ▶ **National Hospital Discharge Data** [<http://www.cdc.gov/nchs/about/major/hdasd/listpubs.htm>]

SECONDARY DATA SOURCES: *National*

- ▶ **National Survey of Children's Health**
[<http://www.nschdata.org/Content/Default.aspx>]
- ▶ **U.S. Census Bureau** [<http://www.census.gov/>]
- ▶ **Department of Education: National Center for Education Statistics**
[<http://nces.ed.gov/>]
- ▶ **U.S. Department of Health & Human Services** [<http://www.hhs.gov>]
 - Administration for Children & families
 - Agency for Healthcare Research and Quality
 - Maternal and Child Health Bureau
 - National Institute of Child Health & Human Development
 - National Institute of Mental Health
 - Substance Abuse & Mental Health Services Administration
 - National Center for Health Statistics
- ▶ **U.S. Department of Housing and Urban Development**
 - Office of Policy Development & Research [<http://www.huduser.org/>]
- ▶ **U.S. Department of Justice**
 - Bureau of Justice Statistics [<http://www.ojp.usdoj.gov/bjs/>]
 - National Institute of Justice [<http://www.ojp.usdoj.gov/nij/>]
 - Office of Juvenile Justice & Delinquency Prevention [<http://ojjdp.ncjrs.org/>]
- ▶ **U.S. Department of Labor**
 - Bureau of Labor Statistics [<http://www.bls.gov/>]
 - Women's Bureau [<http://www.dol.gov/wb/>]
- ▶ **U.S. Department of Transportation**
 - National Highway Traffic Safety Administration [<http://www.nhtsa.dot.gov/>]

SECONDARY DATA SOURCES: *State*

- ▶ **New York State Census Data**

[<http://www.nylovesbiz.com/nysdc/default.asp>]

Census Data provided by the New York State Data Center

- ▶ **New York Census Research Data Center**

[<http://ciser.cornell.edu/NYCRDC/home.shtml>]

The New York Census Research Data Center provides access to Census data in secure facilities.

- ▶ **Vital statistics of New York State**

[http://www.health.state.ny.us/nysdoh/vital_statistics/]

SECONDARY DATA SOURCES: *City*

- ▶ **New York City Department of City Planning**
[<http://www.nyc.gov/html/dcp/html/census/popdiv.shtml>]
Provides demographic statistics about New York City
- ▶ **New York City Department of Health & Mental Hygiene**
[<http://www.nyc.gov/html/doh/html/pr/pr080-05.shtml>]
 - EpiQuery: Epidemiology Interactive Health Data
[<http://www.nyc.gov/html/doh/html/epi/epiquery.shtml>]
 - Community health survey
 - Youth risk behavior survey
 - Vital Statistics
 - Census
 - SPARCS
- ▶ **New York City Mayor's Office of Immigrant Affairs**
[<http://www.nyc.gov/html/imm/html/home/home.shtml>]

SECONDARY DATA SOURCES: *City*

- ▶ **New York City Census FactFinder**

[<http://gis.nyc.gov/dcp/pa/address.jsp>]

The New York City Census FactFinder provides quick and easy access to population information from the 2000 U.S. Census.

- ▶ **New York City Community Boards**

[<http://www.nyc.gov/html/cau/html/cb/directory.shtml>]

- ▶ **Bytes of the Big Apple**

[<http://www.nyc.gov/html/dcp/html/bytes/applbyte.shtml>]

BYTES of the BIG APPLE is a family of software, data and geographic base map files for the City of New York. The Department of City Planning offers these products for free download or by a license agreement.

- ▶ **InfoShare** [<http://www.infoshare.org>]

Profile a local area by factors such as population statistics, immigration trends, socio-economic indicators, birth and death data, hospitalizations, local economic data, and others.

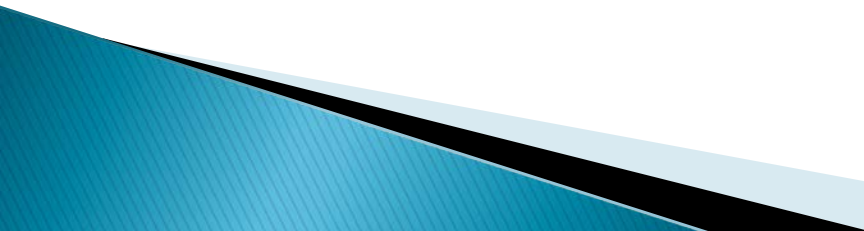
Collecting Quantitative Data

- ▶ Survey research (all require a structured questionnaire, where each observation unit is posed with the exact same questions)
 - mail questionnaire survey
 - e-mail questionnaire survey
 - face-to-face questionnaire survey
 - telephone questionnaire survey
- All methods have limitations

Questionnaire research

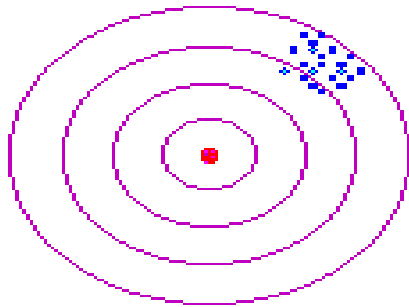
- ▶ Producing worthwhile and generalizable data from questionnaires needs careful planning and design
 - Inappropriate instruments and lack of rigor leads to:
 - Poor quality data
 - Misleading conclusions
 - Vague recommendations
 - Goal: well designed, well managed, and nondiscriminatory questionnaires that contribute to a generalizable evidence

Selecting and designing the questionnaire

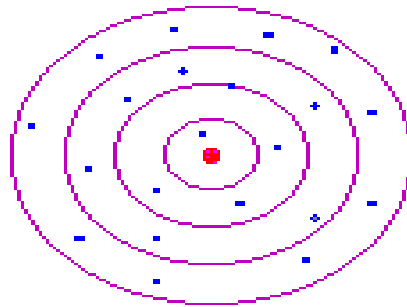
1. What information are you trying to collect?
 - Clarify goals and flag up potential practical problems
 2. Is a questionnaire appropriate?
 - Research participants must be able to give meaningful answers
 3. Could you use an existing instrument?
 - Using a previously validated and published questionnaire will
 - save you time and money
 - allow you to compare your own findings with those from other studies
- 

Selecting and designing the questionnaire (cont.)

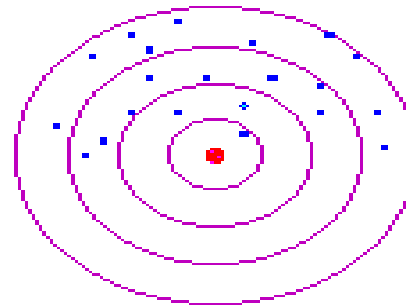
4. Is the questionnaire valid and reliable?
- A valid questionnaire measures what it claims to measure
 - Reliable questionnaires yield consistent results from repeated samples and different researchers



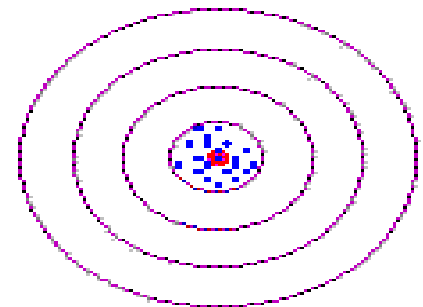
**Reliable
Not Valid**



**Valid
Not Reliable**

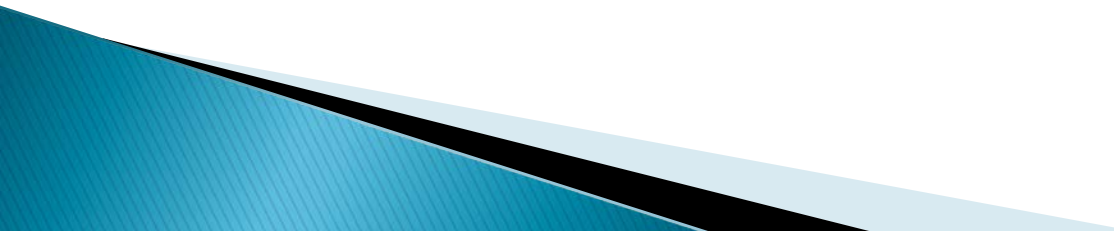


**Neither Reliable
Nor Valid**

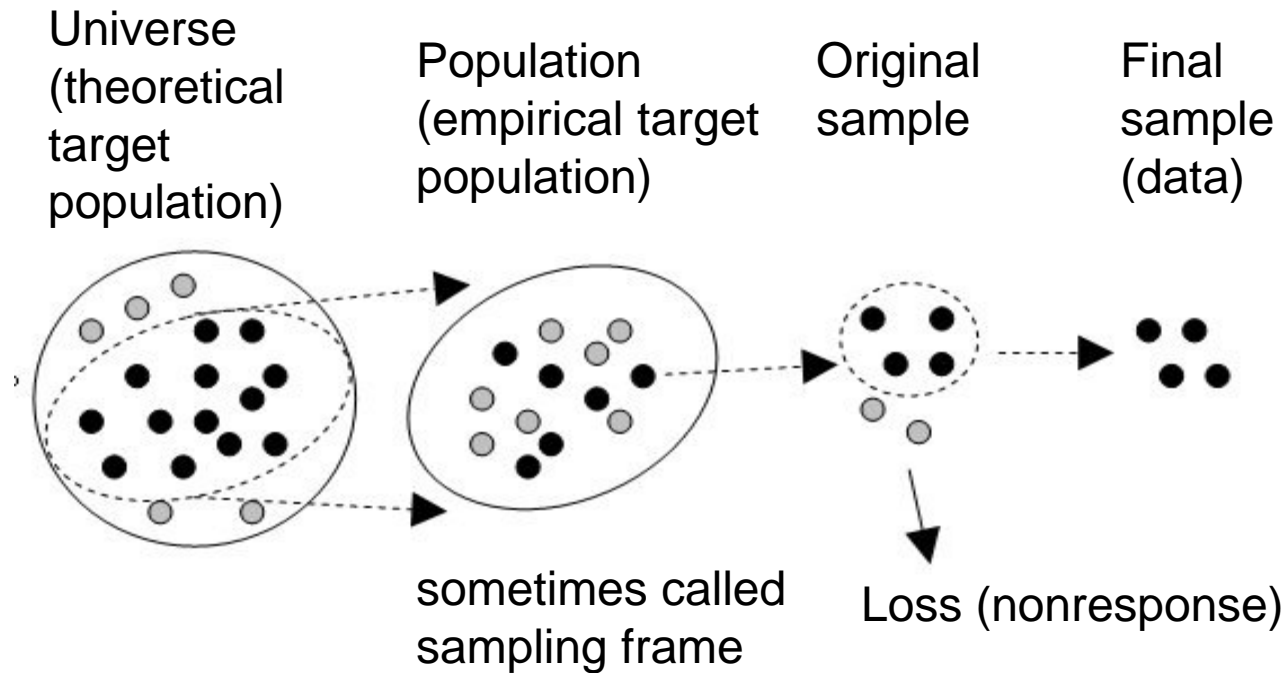


**Both Reliable
And Valid**

Selecting and designing the questionnaire (cont.)

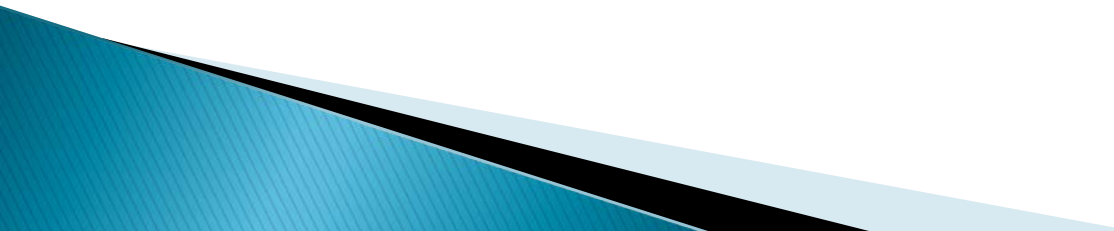
5. How should you present your questions?
 - Items may be opened or closed and be presented in various formats
 6. What should the questionnaire look like?
 - Details such as the font size and color
 - Low response rates are often due to participants being unable to read or follow the questionnaire
 7. How should you select your sample?
 - Different sampling techniques will affect the questions you ask and how you administer your instrument.
- 

Sampling: central concepts



The aim of sampling is to produce a miniature copy of the population. Each member of the population has an equal likelihood of being selected into the sample. Hence we can make inferences about the larger population based on the sample.

Sampling methods

- ▶ Simple random sampling
 - ▶ Systematic selection (interval sampling)
 - ▶ Stratified sampling (proportional and disproportional)
 - ▶ Clustered sampling
 - ▶ Quota sampling (non-probability method)
 - ▶ Convenience sample works with the available participants or other units of analysis
- 

Sampling methods cont.

- ▶ Snowball sample, also known as a respondent – driven sample,



Selecting and designing the questionnaire (cont.)

6. A part from questions, what else should you include?
 - aspects such as introductory letter that should be handed to participants
8. What approvals do you need before you start?

Administering, analyzing, and reporting questionnaires

1. Piloting

- Essential to check the questionnaire works in the study group
- Helps identifying administrative and analytical problems

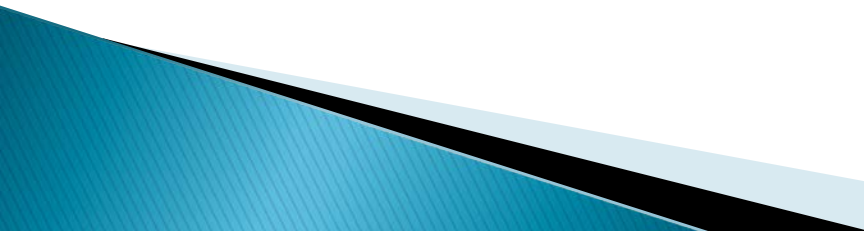
2. Administering the questionnaire

- The method of administration should be determined by scientific considerations (and not just costs!!)

3. Maximizing the response rate

- It is better to collect fewer questionnaires with good quality responses than high numbers that are inaccurate or incomplete

Administering, analyzing, and reporting questionnaires (cont.)

4. Accounting for those who refuse to participate
 - Those who don't participate are equally important
 5. Entering, checking, and cleaning data
 - Should be done as the study progresses
 6. Analyzing data
 - Statistical Analysis: SPSS descriptive, inference
 7. Writing up and reporting
 - Don't try to include all the result when reporting studies
 - Do include the exclusion criteria and data on non-respondents
- 

Qualitative Data Collection Methods

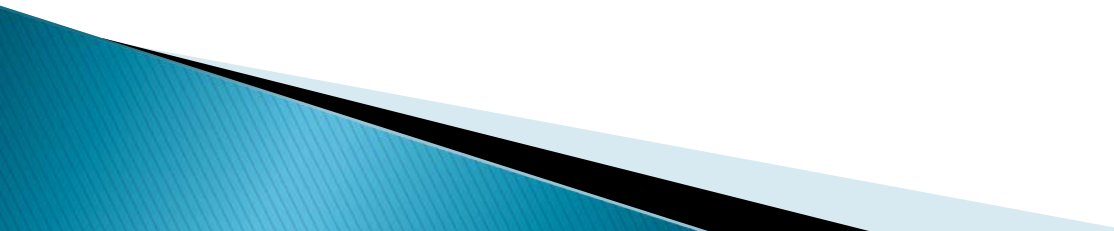
Principles Underlying Qualitative Inquiry

- ▶ Each person or community is unique and deserves respect.
- ▶ Equity, fairness, and mutual respect should be foundations of human interactions.
- ▶ Change, processes (and research) should be negotiated, agreed to, and mutually understood – not imposed, forced, or required.
- ▶ One expresses respect for an concern about others by learning about them, their perspective, and their world – and by being personally involved.
- ▶ Emotion, feeling, and affect are natural, healthy dimensions of human experience.
- ▶ The change agent, therapist, or researcher is nonjudgmental, accepting, and supportive in respecting others' right to make their own decisions and live as they choose. The point is empowerment of others, not control or judgment.
- ▶ People and communities should be understood in context and holistically.
- ▶ The process (how things are done) is as important as the outcome (what is achieved).
- ▶ Action and responsibility are shared; unilateral action is avoided.
- ▶ Information should be openly shared and honestly communicated as a matter of mutual respect and in support of openness as a value.

Sampling methods

- ▶ Sampling methods in qualitative research differs from those used in quantitative research.
- ▶ The sampling numbers are generally smaller than in quantitative research.
- ▶ To get random, representative samples from target population is seldom possible (or the aim) of qualitative research.
- ▶ It is not intended to be statistically representative.
- ▶ Instead, a sample is chosen because it represents *important characteristics* of a population – characteristics that are the main concern in evaluation of research.

Data Collection Methods: Qualitative Data

- ▶ Documents (secondary analysis)
 - ▶ Three data collection strategies introduced:
 1. Participant observation
 2. Interviews (eg Key Informant Interviews)
 3. Focus group interviews
 - ▶ more than one method may be employed
- 

Qualitative Data

▶ Documents:

- Written materials and other documents from organizational, clinical, or programs records;
 - memoranda and correspondence;
 - official publications and reports;
 - personal diaries, letters, artistic works, photographs, and memorabilia
-
- Data consist of excerpts from documents captured in a way that records and preserves context.

Observation Techniques

▶ Description

- Differs from interviewing in that the observer does not actively query the respondent.
- Observation is a method of collecting information about the operation of a natural social process (i.e. communities, political organizations).
- It involves that the researcher
 - Lives or gets involved in the social process studied (**participant observation**)
OR
 - Doesn't try to become a participant in the context but does strive to be as unobtrusive as possible so as not to bias the observations (**direct observer**)

▶ Aim

- A participant observer (in any of such roles) is placed in such natural social process with the aim of collecting more detailed information

▶ Expected outcome

- Information about a social process collected by a participant observer can ensure that it incorporates needs and opinions more acceptable to those who are part of such natural social process

Observation Techniques

- ▶ Participant observer
 - Continuum of roles, ranging from
 - **Complete observer**, who doesn't participate in group activities and is publicly defined as a researcher
 - **Covert participant**, who acts just like other group members and doesn't disclose his or her research role
 - There are middle points in such continuum



Observation Techniques

Strengths	Weaknesses
<ul style="list-style-type: none">■ Researcher is able to experience and therefore to better understand the context of the natural social process and the experiences of the group studied	<ul style="list-style-type: none">■ Presence of the researcher may alter behavior■ Ethical considerations if researcher misrepresents himself/herself (i.e. issues about covert observer)

Key Informant Interviews

- ▶ What are key informant interviews?
 - Qualitative, in-depth interviews of selected people who has first-hand knowledge about the topic of interest
 - Interviews are loosely structure and rely on open-ended questions
 - Rather than asking questions in a fixed order, interviewers allow specific content and order to vary, frame questions spontaneously.
- ▶ When are key informant interviews appropriate?
 - When qualitative information is needed to understand motivations, behaviors, or perspectives
 - When quantitative data previously collected needs to be interpreted
 - When preliminary information is needed to design a future study

Key Informant Interviews

- ▶ Steps in conducting a key informant interview (Kumar, 1986)
 1. Identify concerns of the study
 2. Prepare a short interview guide
 - Key informant interviews do not use rigid questionnaires, however, interviewers need to have an idea of what questions to ask
 3. Select key informants
 - Identify institutions, organizations and groups from which key informants can be identified
 - Select key informant from diverse groups (and with different perspectives)
 - Ask key informants to suggest other people that could be contacted

Key Informant Interviews

4. Conduct Interviews

- Begin with an explanation of the purpose of the interview and assurances of confidentiality
- Start with factual questions, then move to questions that require opinion and judgment
- Begin with questions about the present and then move to the past or future
- Frame questions in a way that leads to answers and explanations (avoid “yes or no” questions)
- Encourage informants to detail basis of their arguments or conclusions
 - KI: “The water program has really changed things around here”
 - I: “What changes have you noticed?”
- Maintain a neutral attitude
- Minimize translation difficulties

5. Take adequate notes

Key Informant Interviews

6. Analyze interview data

- At the end of each interview, prepare a 1–2 pages summary placing information according to themes, issues, key informant position, reason for inclusion of the informant, implications of the points made, etc.
- Coding – systematic recording of data through descriptive codes that help organize information

7. Check for reliability and validity

- Check representativeness if key informants – make sure that no significant groups were overlooked
- Assess reliability of key informants
- Assess interviewer bias (e.g. tendencies to concentrate on information that confirms preconceived notions or hypotheses)
- Get feedback from informants – a summary report can be shared with the informant

Key Informant Interviews

- ▶ Advantages and disadvantages of key informant interviews

Advantages	Disadvantages
<ul style="list-style-type: none">■ Opportunity to get “insiders” view■ Provide flexibility to explore new ideas and issues not anticipated■ Inexpensive and simple to conduct■ Can be combined with other techniques	<ul style="list-style-type: none">■ Not appropriate if quantitative data is needed■ May be biased if informants are not selected carefully■ May overlook perspectives of informants who are less visible■ Susceptible to interviewer biases■ Takes time to select good informants and build trust

Various sources

Focus Groups

- ▶ General aspects:
 - In a focus group the interviewer asks group members specific questions about topics – this requires previous research!
 - Krueger (1988) defines a focus group as a carefully planned discussion designed to obtain perceptions in a defined area of interest.
 - Focus groups are used to collect qualitative data, using open ended questions posed by a researcher or group moderator.
 - Focus groups are usually taped or recorded

Focus Groups

- ▶ **focus groups** – Stewart and Shamdasani (1990) summarize common situations in which focus groups can be considered:
 1. Obtaining general background information
 2. Generating research hypothesis that can be later explored with other types of approaches
 3. Stimulating new ideas and concepts
 4. Diagnosing potential problems (i.e. with a new program)
 5. Generating impressions (i.e. about programs)
 6. Learning how respondents talk about the phenomenon of interest which may facilitate the use of quantitative tools
 7. Interpreting previous results

- ▶ **Size of the groups**
 - Usually 6–12

- ▶ **Participant selection**
 - Participants must be representative of a larger population – convenience sample can be used

Focus Groups

▶ Advantages and disadvantages of Focus Groups

Advantages	Disadvantages
<ul style="list-style-type: none">• Open-ended questions provides opportunity to spontaneously deal with issues as they arise• Cost-effective method of collecting data• Less time-consuming<ul style="list-style-type: none">■ Can be combined with other techniques	<ul style="list-style-type: none">■ Not appropriate if quantitative data is needed■ One or two participants may dominate■ Not done in a natural setting, so little “observation” to help understand the experience of the participants

Collecting Qualitative Data: Managing & Analyzing Qualitative Data

- ▶ Transcribe interview notes and recordings as soon as possible after collecting
- ▶ Qualitative data tends to be voluminous and must be coded by theme to make them more accessible and understandable to others.
- ▶ This can be done by hand or using computer programs such as NUD*IST or ATLAS.ti