Research Design: Methodology

Research design; Methods and tools for data collection and analysis

- Research design addresses the decisions to be made in order to actualize the research. These questions include: what, where, how much, by what means concerning the research.

- "A research design is thus the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure" [Kothari, 2004]

- It forms the conceptual structure within which the research is conducted; the plan for collection, measurement and analysis of data. The goal of sound research design is to provide results that are credible. Credibility here means extent to which results approximate reality and are judged trustworthy and reasonable.

Research design; Methods and tools for data collection and analysis

The research design has the following parts/ components:

- Sampling design ➔ deals with the way of choosing the items to be observed for the given study;
- Observational design ➔ concerns the conditions under which the observations are to be made;
- Statistical design ➔ deals with how many observations are to be made and how the analysis is to be done;
- Operational design ➔ deals with techniques by which the procedures specified in the sampling, statistical and observational designs can be implemented.

Research design; Methods and tools for data collection and analysis

- The important features of a research design are:
  - It's a plan that specifies the sources and types of information relevant to the research problem.
  - It's a strategy giving which approach will be used for gathering and analyzing data.
  - It specifies the resources (financial and time) required to undertake the study.

❖ We thus see that research design is important because it seeks to ensure efficient research operations for maximal information with minimal efforts and money expended.

Research design; Methods and tools for data collection and analysis

- Research design = advance planning for the methods of collection of relevant data, analysis, and interpretation observations. It is the logical model of proof that allows the researcher to draw inferences concerning causal relationships among the variables under investigation

- It forms the foundation upon which reliability (Quantitative)/dependability (Qualitative) of the research results are founded

- Hence the research design forces the researcher to organize his/her ideas in a manner that reveals flaws and inadequacies.

Components of Research Designs

- The four components are:

1. Comparison – an operation req'd to demonstrate the 2 variables are correlated, thus it underlies the concept of covariation.

2. Manipulation – the operation that controls the assignment to the treatment group, so the researcher can decide the time sequence to make sure the independent variable changes before the dependent variable.

3. Control – the operation that enables the researcher to rule out the rival explanations for the change of the dependent variable.

4. Generalizability – the extent to which the research findings can be generalized to larger populations and applied to different setting.

Research design; Methods and tools for data collection and analysis

- A good design is flexible, appropriate, efficient and economical

The following concepts are important in research design

- Dependent and independent variables:
  - A variable is that concept/ anything that can assume different quantitative values
  - A variable can be discrete (takes on only specific values, e.g. number of children) or continuous (takes on any values, e.g. age)
- Dependent variable: that which depends on or is a consequence of another variable (e.g. height is dependent on age)--→Used a lot in tradionalist Kenya for students to start school

- Independent variable: the antecedent of the dependent variable, i.e., what the experimenter, changes or enacts in order to the experiment

Research design; Methods and tools for data collection and analysis

- Extraneous variable: an independent variable not related to the purpose of the study but may affect the dependent variable

- Control: This term is used to mean designing a study to minimize the effects of extraneous variable(s), i.e., is used to refer to the restraints on the experimental conditions

- Experimental and control groups: The group exposed to the usual conditions is called the "control group" but the group exposed to the special conditions or treatments is called the "experimental group"

The roles of variables change with situations, and also depend the theoretical model advanced in relating the variables

Research design; Methods and tools for data collection and analysis

- Treatments: the different conditions to which the experimental and control groups are exposed

- Experiment: The process of examining the truth of the problem of the study with the view to generating the desired information. The experiment can either be absolute (e.g. comparing the impact of fertilizer on crop yield) or comparative (e.g. comparing the efficacies of two types of fertilizer on crop yield)

- Sources of Variability in design include: systematic, error and extraneous. Systematic variance is related to the variables to be investigated (hence needs to maximized in the design).
  - Error variance need to be minimized.
  - Extraneous variance needs to be controlled.

Research design; Methods and tools for data collection and analysis

- We have already noted that that research design is a plan/strategy of how the problem of the study will be investigated.

- This strategy should ensure:
  - ❑ Validity (the internal and external representation of the results) is maintained, i.e., measuring was intended to be measured

    - ○ Internal validity ascertains that the results obtained are attributed to the treatment variable. It is thus a measure of the internal reliability of the study. This calls for careful control of the internal (e.g., changes in history, maturation, exp'tal mortality, instrumentation, etc) and external (non-random assignment of subjects) influences (normally called threats) that may interfere or affect the variable being studied. These threats include: Factors that affect data collection methods (e.g. extended time, fatigue etc.); and use of untrained research assistants.

Research design; Methods and tools for data collection and analysis

- External validity ensures that the results obtained would apply to or represent similar populations, i.e., it refers to the generalizability of the results. The two issues of external validity are: the representativeness of the sample and reactive arrangements in the research procedure (i.e., a study carried out in a highly artificial environment).

  - This requires that the sample selected is a representation of the population. Hence correct sampling procedures (to avoid bias) and size of sample is to be followed to ensure external validity.

- Reliability of measure: This is a requirement that the same results will be obtained by another experimenter under similar conditions on the same samples

Research design types

In your small proposal writing groups, familiarize yourselves with the different **research designs**. **Read through the course book (Kothari & Kothari) from page (35-52) for a better description.**

At the end of the read, you should all start having an idea of which research design to use as you write your chapter 3

Class Discussion (1Hour)

Sampling design

- Sampling theory: A correctly taken sample of an appropriate size will yield results that can be applied to the population as a whole.

- A sample is part of the population chosen for study ($n < N$ where n is the sample size and N the population size).

- It is the practice to sample in order to reduce cost in terms of money, time and energy.

- It is also important that any bias in the inquiry gets larger with increase in the number of observations, and this can only be corrected by another expensive survey.

- The sample size, n, which part of the larger population is selected according to some rules ➔ sample design.

Sampling design

- A sample design thus is a plan of obtaining a sample from the population

- It is the technique/procedure adopted in selecting the items to included in the sample

The steps to follow in sample design

- Define the items in the Universe (which can be finite or infinite)
- Decide the sampling unit
- Draw the source list or sampling frame (which must be comprehensive, correct, reliable and appropriate)
- Determine the size of the sample (i.e., no. of items from the population to chosen) (* a major problem for the researcher)

Sampling design

** The sample size should be optimum (that meets the requirements of efficiency, representativeness, reliability and flexibility)

- Bare in mind the parameters of interest (i.e., research context)
- Consider the prevailing budgetary constraints
- Decide the sampling procedure.

Remember

- The theory of sampling is based on random samples – where all items in the population have the same chance of being selected as sample units.

Sampling design

- Three types of random sample can be drawn – a simple random sample (SRS), a stratified sample and a systematic sample.

- Simple random sampling can be carried out in two ways – the lottery method and using random numbers.

- It has two disadvantages: (i) A sampling frame is required and this may not be available, exist or be incomplete; (ii) The procedure is unbiased but the sample may be biased, e.g., if the population is not homogeneous, e.g., a mixture of men and women (population stratified) and all men were selected this would be a biased sample.

- Sample bias is overcome by taking a stratified sample so that the stratified population structure is reflected in the sample structure, subject to some criterion

Sampling design

- Systematic sampling involves choosing the every *ith* item from the population

- Though not truly random this is a method popular because it is easy to operate and quick, even when the population and the sample are large

Methods and tools for data collection and analysis

- It's always important the researcher collects data in a form that s/he can understand or analyse.

- When collecting data, ensure: reliability (extent to which the same finding will be obtained if the research was repeated at another time by another researcher); Validity (measuring what was intended); Triangulation (crosschecking of data using multiple data sources or using two or more methods of data collection - helps check researcher's biases).

- Data can be obtained at two levels: primary data (the researcher collects the data), and secondary data (researcher uses existing data)

Data collection cont'd

- Effective and efficient data collection requires careful planning and judicious sue of the researcher's time
- For example, in a case study, have sufficient background information about the study site before starting data collection
- Organize and properly categorize the data as this will facilitate the task of data analysis
- Plan ahead for dealing with difficulties of gathering data that may arise, especially in real-life environment not controlled by the researcher
- Develop a general data analysis strategy as part of the design as this will ensure that data collection activities are appropriate and support the ways in which evidence is to be analyzed

Methods and tools for data collection and analysis

- Methods of primary data collection include:
  - Questionnaires
  - Interviews
  - focus group interviews
  - Observation/experimentation
  - case-studies
  - Diaries
  - critical incidents
  - Portfolios

  (Read details of each method of data collection)

- Secondary data is data that has already been collected by someone else for a different purpose to the researcher's

Observation

- Considered the standard method of scientific research as it aids understanding, explanation of an object of study, and any predictions
- Suitable for investigating phenomena that can be observed directly
- For systematic findings, the observations must be done with reference to the following crucial issues: (i) what to observe, (ii) where and when to observe, (iii) how to infer when recording observations
- How to carry out observations is the focus of the research problem and research design
- When the aim of the researcher is to test a hypothesis experimentally, the units of observations are explicitly defined, a setting chosen (laboratory or field), a time sample drawn, and observations systematically recorded – typical processes of a controlled observations

Survey research

- Suitable where the phenomena may not be accessible to the researcher to observe directly

- Survey research may comprise use of (i) mail questionnaire, (ii) personal interview, (iii) telephone interview (survey)

- Mail questionnaire:

  Advantages – low cost, reduces biasing error that may result from personal characteristics of the interviewers or variability in their skills, greater anonymity, considered answers and consultations, provides wider access

  Disadvantages – requires simple questions, no opportunity for probing for additional information, no control on who fills out the questionnaire, low response rate

Personal interview

- A face-to-face, interpersonal role played by the researcher asking respondents questions designed to elicit answers pertinent to the research problem

- The questions, their wordings, and sequence define the structure of the interview. Types include:

  The schedule-structured interview – where the number of questions and their wordings are identical to all respondents (this very rigid);

  The focused interview (non-schedule-structured) – has four characteristics: respondents have been involved with a particular experience, situations have analyzed prior to interview, there is an interview guide specifying the topics related to the hypotheses, focused on the respondents' experience on the situation under study

  The nondirective (nonstructured) interview – most flexible as respondent is not subjected to prespecified set of questions or questions asked in a specified order

Personal interview

- Advantages include: Flexibility in questioning process; control of interview situation; High response rate; Fuller information (supplementary questions possible)

- Disadvantages include: Higher cost, especially for spatially dispersed respondents; Interviewer bias; Lack of anonymity

Principles of Interviewing

1. The respondents need to feel that their interaction with the interviewer will be pleasant and satisfying

2. The respondents need to see the study as being worthwhile

3. Barriers to the interview in the respondents' minds need to be overcome

Telephone Interview (survey)

- A semi-personal interview.
- Advantages include: Moderate cost; speed; high response rate; high quality data may be collected when the interviewers are centrally located and closely supervised

- Disadvantages include: Reluctance to discuss sensitive topics; Termination of interview before it is complete by respondents; Less information (i.e., information on the respondents' characteristics or environment)

Questionnaire Construction

- Questionnaires are key in survey research. The foundation of questionnaires is the Question.
- The questionnaire must translate the research objectives into specific questions
- Responses to these questions give the data for hypotheses testing
- The major consideration in formulating questions are their content, structure, format and sequence
- The content is concerned with facts, opinions, respondents' motivation, expertise/knowledge of a given subject. Generally, questions can either factual or subjective experiences.

Questionnaire Construction

Structures (Types) of questions include:

- Close-Ended Questions: The respondent is offered a set of answers and asked to choose those that closest to their views

- They are easy to ask, require no writing either by the researcher or respondent; easy analysis, however may introduce bias

- Open-Ended Questions – not followed by any specified choice, therefore respondents' answers are recorded in full

- Have the advantage that the respondents' have latitude for response

- Open-ended questions are difficult to answer, and also to analyze

Questionnaire Construction

Sequence of questions – this is the order in which the questions are placed in the questionnaire

- Two patterns are found appropriate: - Funnel sequence and the inverted funnel sequence

Funnel sequence – each successive question is related to the previous one, and of progressively narrower scope

- Useful when the objective of the survey is obtain full information

- Inverted funnel sequence – narrower questions are followed by broader ones

Questionnaire Construction – Pitfalls to Watch out for

- Wording – questions must be worded in a manner understandable by the respondents and in similar manner
- Response set – tendency to answer all questions in a specific direction regardless of their content. This is caused by presenting a set of questions together with the same response format. This is avoided by changing the question format either by varying the response categories for each question or distributing questions on atopic throughout the questionnaire
- Leading questions – question(s) framed in a manner as to suggest the researcher expects a certain answer. These should be avoided
- Threatening questions – these are anxiety-arousing questions to the respondent. They lead to bias. (Avoid by using open-ended questions)
- Instructions – should be included with any questions that are not self-explanatory
- Double-barreled questions – Questions that combine two or more questions in one
- Cover Letter – Introductory letter to the respondent

Measurements

- Measurement is a procedure of assigning (= mapping) numerals (e.g., I, II, III, … or 1, 2, 3, …) (numbers) or other symbols to the empirical properties (variables) according to rules (= procedures defined) (Nachmias & Nachmias, 1996). It's a complex task especially for abstract or qualitative phenomena (e.g. motivation)

- The numerals (end product of measurement) can be used for comparing, evaluating and assessing relations between the various properties measured

- Note that numerals have no quantitative meanings unless given (implying numerals can also be used to identify objects or phenomena)

- Numerals with quantitative meanings become numbers (thus enable mathematical or statistical analysis)

- Rules are the most significant component of measurement as they determine the quality of measurement (e.g., measurement not tied to reality lack empirical basis and is therefore meaningless)

Level of Measurements

They include:

1. Nominal level – the lowest level where numbers or other symbols are used to classify objects or observations into a number of categories (e.g., 1 for males and 2 for females). At this level, the basic property is that of equivalence, i.e., the properties of objects in one category are designated as identical for all its cases. Nominal data are numerical in name only since they do not share the properties of the numbers as used in arithmetic, e.g., as used above, we cannot perform the operation 4 +2 = 6.

2. Ordinal level – here the variables exhibit some relation (designated by the "greater than" or inequalities symbols, e.g., > (higher than), etc) to each other. This level also has the property of equivalence. For example, measurement of hardness in Mohs, where values such as 5, 6, 7,etc can only be said that 7>6, 6>5 etc. The inequality can also be used to mean "preferred to" etc.

Level of Measurements

3. Interval level – here the exact distance btwn each observation is known and constant. Several arithmetic operations can be performed. Phenomena has property of equivalence. E.g., beside inequalities, we can perform differences. E.g., consider temperature data such as 10˚C, 20˚C, 30˚C, 40˚C, 50˚C, etc. We can say 20˚C > 10 ˚C, 30˚C > 20˚C etc but also perform 20˚C - 10˚C, etc simply because we consider an interval.

4. Ratio level – used to measure variables that have natural zero points (e.g., temperature, time, area, etc). The level also has properties of equivalence, and greater than.

Levels of measurements and their characteristic properties

Level	Equivalence	Greater than	Fixed interval	Natural zero
Nominal	Yes	No	No	No
Ordinal	Yes	Yes	No	No
Interval	Yes	Yes	Yes	No
Ratio	Yes	Yes	Yes	Yes

Measurement and Scaling

- The distinction between nominal, ordinal, interval and ratio data is important for the nature of set of data and may even suggest the type of particular set of statistical techniques to be used in analysis.

Data Analysis

- Data once collected must be processed and analyzed to draw inferences & comparisons
- The processing and analysis is in accordance with the what was put down for the purpose during the development of the research plan (e.g., what to analyze and why)
- Processing includes such actions such as: editing (examining raw data from e.g. surveys to detect error, omissions etc.,), coding, classification and tabulation
- Analysis involves computation of certain measures with a view to finding patterns of relations between the variables, and subjecting the same to statistical tests to determine significance to determine the validity of the 'conclusion'
- Data analysis focuses on discovering regularities and patterns within the data

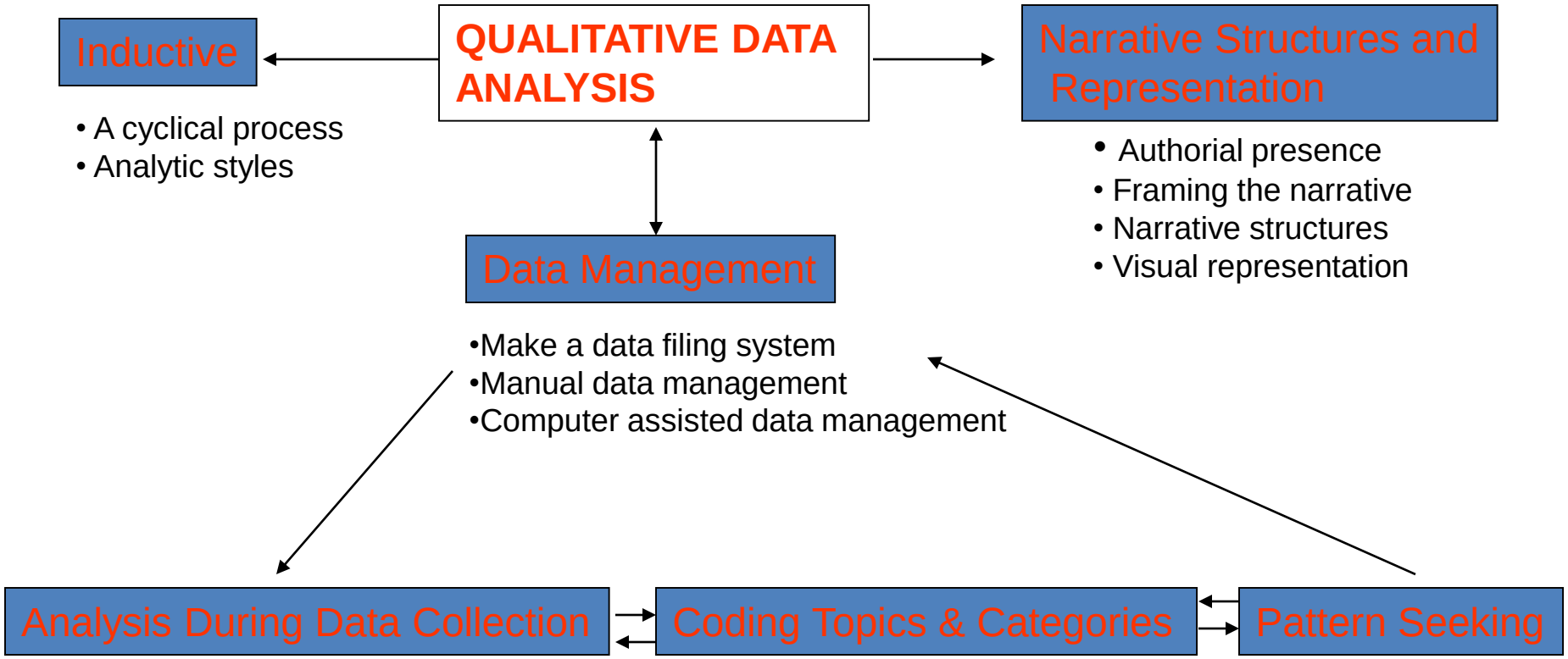
Data Analysis

- Data analysis consists of three concurrent activities

1. Data reduction – process of selecting, simplifying, abstracting and transforming the raw data
2. Data display – organized assembly of information to facilitate the drawing of conclusions
3. Conclusion drawing/verification – drawing meaning from data and building logical chain of evidence

NB Data collection and data analysis should overlap to allow flexibility in data collection procedures and also this makes the researcher alert to new ideas or data which may emerge.

Qualitative Data Analysis



- Discovery analysis strategies
  - observer comments
  - idea playing
  - literature explored
  - metaphors & analogies
- Interim analysis

- Develop an organizing system
- Code segments for topics
- Code categories for meaning

- Techniques
  - data trustworthiness
  - data triangulation
  - discrepant/negative evidence
  - category ordering & sorting
  - visual representations
  - logical cross-analysis
- Plausible explanations

Source: J.H. McMillan & S. Schumacher, Research in Education, A conceptual Introduction 5th Edition

Data analysis - Coding

- Defn: Coding is the assigning of a number to an observation or grouping various classifications of a concept or it is the process by which responses are classified into meaningful categories

- A code should be consistent across cases or units of analysis when the same conditions exist

- Coding schemes: Systems used to categorize /classify responses to a single item or variable. The research questions, theory and knowledge of the subject matter and the sample determine the detail of categories in the coding scheme

- Coding Rules: (i) the numbers assigned must make intuitive sense (ii) the coding should consider theory (especially for deductive coding – common with quantitative research) from which the research hypotheses/questions are derived, mutual exclusivity (each response falls into one and only one category of the variable), and exhaustiveness (every response falls into a category with few being classified as "other") (Inductive coding used mainly in qualitative research) (iii) the categories must be specific enough to be differentiable using the least number of categories (iv) keep the coding scheme simple as this increases coding reliability

Data Analysis, Interpretation

- Data once analyzed needs to be interpretated
- What do the results suggest?
- Has the problem been answered by the results?
- Do we now have a better understanding of the problem compared to the previous studies?
- Interpretation requires deep understanding or knowledge of the problem area (recall literature review)

Data Analysis, Interpretation

Technique of Interpretation often involves:

- Practice & experience, and seeking guidance from experts
- Giving explanations to the relations which the research has found in line with the underlying processes, and trying to find out any thread of uniformity. This aids generalizations
- Considering also any extraneous information collected during the study
- Discuss/consult with experts
- Consider all the relevant factors affecting the problem in order to avoid false generalizations, i.e., do not rush

Data Analysis, Interpretation

Pay attention to:

- Appropriateness and reliability of data and proper analysis
- Possible sources of errors that can arise during the interpretation, especially from false generalization and incorrect use of statistical measures
- The close interlinkage between analysis and interpretation and the two cannot be separated hence make computational checks, validation and comparison of results
- Peculiar observations, occurrences, and unmask any hidden factors
- The interconnections and interactions between the empirical observations and theoretical conceptions. This provides the opportunities for originality and creativity.

Some references

- Chen, W. S. and Hirschheim (2004) A paradigmatic and methodological examination of information systems research from 1991 to 2001. *Information Systems Journal,* **14**, 197 – 235

- Darke, P., Shank, G., and Broadbent, M. (1998) Successfully completing case study research: combining rigour, relevance and pragmatism. *Information Systems Journal*, **8**, 273 -289

- Plummer, A.A. (2001) Information systems methodology for building theory in health informatics: the argument for a structured approach to case study research. Proceedings of the 34th Hawaii International Conference on System Sciences, pp 1 -10

- Halonen, R. (2009) Practice-related research methods as part of scientific information systems research. *Indian Journal of Science and Technology* 2(5) 75 - 81