

Addis College

M.Sc : Construction Technology & Management

Course: Research Methods and Statistical Application

Lecture 5: Research Design

By: Dagnachew A. (PhD)

Contents of the presentation

- Lecture notes
- Individual assignment

Research Design

- To define a research problem the preparation of the design of the research is needed.
- Decisions regarding what, where, when, how much, by what means concerning an inquiry or a research study constitute a research design.
- It is the arrangement of conditions for collection and analysis of data.
- It is the conceptual structure constitutes the blueprint for the collection, measurement and analysis of data.
- The design includes an outline of what the researcher will do from writing the hypothesis and its operational implications to the final analysis of data. More explicitly, the design decisions should be done in respect of:

I. What is the study about?

II. Why is the study being made?

III. Where will the study be carried out?

IV. What type of data is required?

V. Where can the required data be found?

VI. What periods of time will the study include?

VII. What will be the sample design?

VIII. What techniques of data collection will be used?

IX. How will the data be analyzed & presented?

X. In what style will the report be prepared?

The overall research design might be divided into the following parts:

- a. the sampling design which deals with the method of selecting items to be observed for the given study;*
- b. the observational design which relates to the conditions under which the observations are to be made;*
- c. the statistical design which concerns with the question of how many items are to be observed and how the information and data gathered are to be analyzed & presented; and*
- d. the operational design which deals with the techniques by which the procedures specified in the sampling, statistical and observational designs can be carried out.*

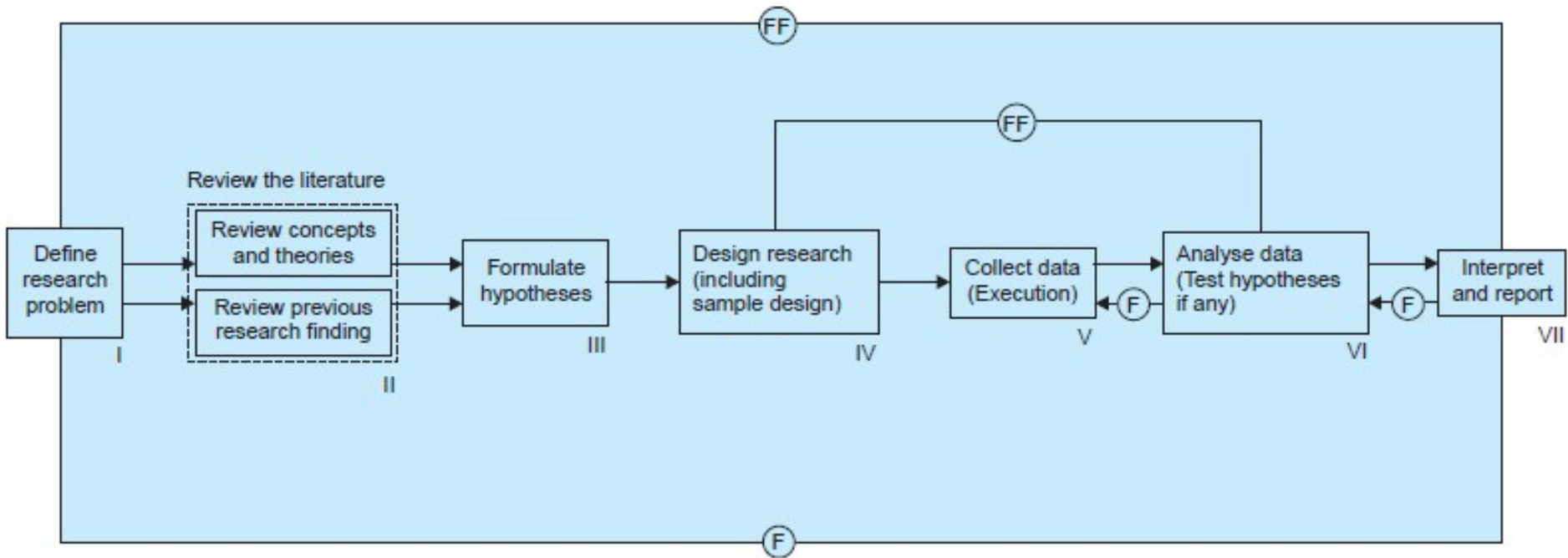
Important features of a research design

- I. It is a plan that specifies the sources and types of information relevant to the research problem.
- II. It is a strategy specifying which approach will be used for gathering and analyzing the data.
- III. It also includes the time and cost budgets since most studies are done under these two constraints.

Generally, research design must, at least, contain:

- (a) a clear statement of the research problem;
- (b) procedures and techniques to be used for gathering information;
- (c) the population to be studied; and
- (d) methods to be used in processing, analyzing data & presenting data.

Research Process



Where (F) = feed back (Helps in controlling the sub-system to which it is transmitted)
(FF) = feed forward (Serves the vital function of providing criteria for evaluation)

Formulating the research problem:

There are two types of research problems:

- At the very beginning the researcher must single out the problem he wants to study, i.e., s/he must decide the general area of interest or aspect of a subject-matter that he would like to inquire into.
- Initially the problem may be stated in a broad general way and then the ambiguities, if any, relating to the problem be resolved.
- Then, the feasibility of a particular solution has to be considered before a working formulation of the problem can be set up.
- The formulation of a general topic into a specific research problem, thus, constitutes the first step in a scientific enquiry.
- Essentially two steps are involved in formulating the research problem:
 1. understanding the problem thoroughly, and
 2. rephrasing the same into meaningful terms from an analytical point of view.

- The researcher must at the same time examine all available literature to get him/herself acquainted with the selected problem.
- S/he may review **two types of literature** - the *conceptual literature* concerning the concepts and theories, and the *empirical literature* consisting of studies made earlier which are similar to the one proposed.
- The basic outcome of this review will be the knowledge as to what data & other materials are available for operational purposes which will enable the researcher to specify his/her own research problem in a meaningful context.
- After this the researcher rephrases the problem into analytical or operational terms i.e., to put the problem in as specific terms as possible.

- This task of formulating, or defining, a research problem is a step of greatest importance in the entire research process.
- The problem to be investigated must be defined unambiguously for that will help discriminating relevant data from irrelevant ones.
- Care must, however, be taken to verify the objectivity and validity of the background facts concerning the problem.

Extensive literature survey

Once the problem is formulated, a brief summary of it should be written down.

An extensive literature survey should be connected with the research problem.

For this purpose, a number of sources including the abstracting & indexing journals and published or unpublished bibliographies are the first place to be reviewed. Academic journals, conference proceedings, government reports, books etc., are the other sources to be used.

In this process, it should be remembered that one source will lead to another.

Earlier studies, if any, which are similar to the study in hand should be carefully reviewed.

A good library will be a great help to the researcher at this stage

Development of working hypotheses

- After extensive literature survey, researcher should state the working hypothesis or hypotheses.
- Working hypothesis is tentative assumption made to draw out and test its logical or empirical consequences.
- As such the manner in which research hypotheses are developed is particularly important since they provide the focal point for research.
- They also affect the manner in which tests must be conducted in the analysis of data and indirectly the quality of data which is required for the analysis. In most types of research, the development of working hypothesis plays an important role.
- Hypothesis should be very specific and limited to the piece of research in hand because it has to be tested.
- The role of the hypothesis is to guide the researcher by delimiting the area of research and to keep him/her on the right track.

- ***Techniques to develop working Hypothesis:***

- a. Discussions with colleagues and experts about the problem, its origin and the objectives in seeking a solution;
- b. Examination of data and records, if available, concerning the problem for possible trends, peculiarities and other clues;
- c. Review of similar studies in the area or of the studies on similar problems; and
- d. Exploratory personal investigation which involves original field interviews on a limited scale with interested parties and individuals with a view to secure greater insight into the practical aspects of the problem.

Determining sample design:

All the items under consideration in any field of inquiry constitute a 'universe' or 'population'.

A complete enumeration of all the items in the 'population' is known as a census inquiry.

It can be presumed that in such an inquiry when all the items are covered no element of chance is left and highest accuracy is obtained. But in practice this may not be true.

Even the slightest element of bias in such an inquiry will get larger and larger as the number of observations increases.

Moreover, there is no way of checking the element of bias or its extent except through a resurvey or use of sample checks.

Besides, this type of inquiry involves a great deal of time, money and energy. Not only this, census inquiry is not possible in practice under many circumstances.

We select only a few items from the universe for our study purposes. The items so selected constitute what is technically called a *sample*.

- A sample design is a definite plan determined before any data are actually collected for obtaining a sample from a given population.
- For example, the plan to select 12 of a city's 200 construction industries in a certain way constitutes a sample design.
- Samples can be either *probability* samples or *non-probability* samples. With **probability** samples each element has a known probability of being included in the sample but the **non-probability** samples do not allow the researcher to determine this probability.
- Probability samples are those based on simple **random** sampling, **systematic** sampling, **stratified** sampling, cluster/area sampling whereas non-probability samples are those based on convenience sampling, judgment sampling and quota sampling techniques.

Types of sampling

1. **Purposive/Deliberate sampling**

- *Deliberate sampling is also known as purposive or non-probability sampling.*
- It involves purposive or deliberate selection of particular units of the universe for constituting a sample which represents the universe.
- When population elements are selected for inclusion in the sample based on the ease of access, it can be called *purposive sampling*.
- *If a researcher wishes to secure data from, say, construction materials buyers, the researcher may select a fixed number of construction materials shops and may conduct interviews at these shops.*
- Sometimes such a procedure may give very biased results particularly when the population is not homogeneous.
- On the other hand, in *judgment sampling the researcher's judgment is used* for selecting items which he considers as representative of the population.
- For example, a judgment sample of college students might be taken to secure reactions to a new method of teaching.
- Judgment sampling is used quite frequently in qualitative research where the desire happens to be to develop hypotheses rather than to generalize to larger populations.

2. **Simple random sampling**

- *This type of sampling is also known as chance sampling or probability sampling where each and every item in the population has an equal chance of inclusion in the sample and each one of the possible samples, in case of finite universe, has the same probability of being selected.*
- For example, if we have to select a sample of 300 items from a universe of 15,000 items, then we can put the names or numbers of all the 15,000 items on slips of paper and conduct a lottery.
- To do this we select some random starting point and then a systematic pattern is used in proceeding through the table.
- When a number exceeds the limit of the numbers in the frame, in our case over 15,000, it is simply passed over and the next number selected that does fall within the relevant range.

3. **Systematic sampling**

- *In some instances the most practical way of sampling is to select every 15th name on a list, every 10th house on one side of a street and so on. Sampling of this type is known as systematic sampling.*
- An element of randomness is usually introduced into this kind of sampling by using random numbers to pick up the unit with which to start.
- This procedure is useful when sampling frame is available in the form of a list.
- In such a design the selection process starts by picking some random point in the list and then every n^{th} element is selected until the desired number is secured.

4. *Stratified sampling*

- *If the population from which a sample is to be drawn does not constitute a homogeneous group, then stratified sampling technique is applied so as to obtain a representative sample.*
- In this technique, the population is stratified into a number of non-overlapping subpopulations or strata and sample items are selected from each stratum.
- If the items selected from each stratum is based on simple random sampling the entire procedure, first stratification and then simple random sampling, is known as *stratified random sampling*.

5. Quota sampling

- *In stratified sampling the cost of taking random samples from individual strata is often so expensive that interviewers are simply given quota to be filled from different strata, the actual selection of items for sample being left to the interviewer's judgment. This is called quota sampling.*
- The size of the quota for each stratum is generally proportionate to the size of that stratum in the population.
- Quota sampling is thus an important form of non-probability sampling. Quota samples generally happen to be judgment samples rather than random samples.

6. Cluster sampling

- *It involves grouping the population and then selecting the groups or the clusters rather than individual elements for inclusion in the sample.*
- Suppose some departmental store wishes to sample its credit card holders. It has issued its cards to 15,000 customers.
- The sample size is to be kept say 450. For cluster sampling this list of 15,000 card holders could be formed into 100 clusters of 150 card holders each. Three clusters might then be selected for the sample randomly.
- The sample size must often be larger than the simple random sample to ensure the same level of accuracy because in cluster sampling procedural potential for order bias and other sources of error is usually accentuated.
- The clustering approach can, however, make the sampling procedure relatively easier and increase the efficiency of field work, specially in the case of personal interviews.

Data collection

- There are several ways of collecting the appropriate data which differ considerably in context of costs, time and other resources at the disposal of the researcher.
- Primary data can be collected either through *experiment or through survey*.
- If the researcher conducts an experiment, s/he observes some quantitative measurements, or the data, with the help of which s/he examines the truth contained in his hypothesis.

In the case of a survey, data can be collected by any one or more of the following ways:

- **Observation:** *This method implies the collection of data by way of investigator's own observation, without interviewing the respondents. The data obtained relates to what is currently happening & is not complicated by either the past behavior or future intentions or attitudes of respondents. This method is no doubt an expensive method and the information provided by this method is also very limited. This method is not suitable in inquiries where large samples are concerned.*
- **Personal interview:** *The investigator follows a rigid procedure and seeks answers to a set of pre-conceived questions through personal interviews. This method of collecting data is usually carried out in a structured way where output depends upon the ability of the interviewer to a large extent.*

Questionnaires: *The researcher and the respondents do come in contact with each other if this method of survey is adopted.*

- Questionnaires are distributed in person or mailed to the respondents with a request to return after completing the same.
- It is the most extensively used method in various economic and business surveys.
- Before applying this method, usually a Pilot Study for testing the questionnaire is conducted which reveals the weaknesses, if any, of the questionnaire.
- Questionnaire to be used must be prepared very carefully so that it may prove to be effective in collecting the relevant information.

Telephone interviews: *This method of collecting information involves contacting the respondents on telephone itself.*

- This is not a very widely used method but it plays an important role in industrial surveys in developed regions, particularly, when the survey has to be accomplished in a very limited time.

- *The researcher should select one of these methods to collect the required data taking into consideration:*
 - ✓ *the nature of investigation,*
 - ✓ *objective and scope of the inquiry,*
 - ✓ *financial resources,*
 - ✓ *available time and the desired degree of accuracy.*
- *Such data collection depends upon the ability and experience of the researcher.*

Execution of the Research

- Execution of the research is a very important step in the research process.
- If the execution of the research proceeds correctly, the data to be collected would be adequate and dependable.
- If the survey is to be conducted by means of structured questionnaires, data can be readily machine-processed. In such a situation, questions as well as the possible answers may be coded.
- If the data are to be collected through interviewers, arrangements should be made for proper selection and training of the interviewers.
- Occasional field checks should be made to ensure that the interviewers are doing their assigned job sincerely and efficiently.
- A careful watch should be kept for unanticipated factors in order to keep the survey as much realistic as possible.
- This means that steps should be taken to ensure that the survey is under statistical control so that the collected information is in accordance with the pre-defined standard of accuracy.

Analysis of data

- After the data have been collected, the researcher turns to the task of analyzing them.
- The analysis of data requires a number of closely related operations such as establishment of categories, the application of these categories to raw data through coding, tabulation and then drawing statistical inferences.
- The bulky data should necessarily be condensed into a few manageable groups and tables for further analysis. Thus, researcher should classify the raw data into some purposeful and usable categories.
- **Coding** *is usually done at this stage through which the categories of data are transformed into symbols that may be tabulated and counted.*
- **Editing** *is the procedure that improves the quality of the data for coding. With coding the stage is ready for tabulation.*
- **Tabulation** *is a part of the technical procedure wherein the classified data are put in the form of tables. Computers can be used large inquiries.*
- Computers not only save time but also make it possible to study large number of variables affecting a problem simultaneously.

- For instance, if there are two samples of weekly wages, each sample being drawn from factories in different parts of the same city, giving two different mean values, then our problem may be whether the two mean values are *significantly different* or the difference is just a *matter of chance*.
- Through the use of statistical tests we can establish whether such a difference is a real one or is the result of random fluctuations. If the difference happens to be real, the inference will be that the two samples come from different universes and if the difference is due to chance, the conclusion would be that the two samples belong to the same universe.
- Similarly, the technique of analysis of *variance* can help us in analyzing whether three or more varieties of cements yield significantly different results or not.
- In brief, the researcher can analyze the collected data with the help of various statistical measures.

Data Presentation

- After analysis the data should be presented using appropriate statistical tools such as *graphs, charts (e.g. pie chart,...), percentages, coefficients*, etc., by applying various well defined statistical formula.

Hypothesis-testing

- After analyzing the data as stated above, the researcher is in a position to test the hypotheses, if any, he had formulated earlier.
- Do the facts support the hypotheses or they happen to be contrary? This is the usual question which should be answered while testing hypotheses.
- Various statistical methods can be used such as **Chi square test, t-test, F-test**, *can be developed by statisticians or the researcher him/herself for the purpose.*

- The hypotheses may be tested through the use of one or more of such tests, depending upon the nature and object of research inquiry.
- Hypothesis-testing will result in either accepting the hypothesis or in rejecting it. If the researcher had no hypotheses to start with, generalizations established on the basis of data may be stated as hypotheses to be tested by subsequent researches in times to come.

Discussion/Interpretation of Results

- After the data/results are presented in line with the order of the specific objectives, the implications of the results should be interpreted/discussed.
- Moreover, the present results should be compared with previous (related) studies, whether they are consistent or not. If not, the reasons should be presented.

Statistical methods/Tests

Some common statistical methods used in data analysis are :

1. non-parametric tests

- sign test
- rank-sum test
- *chi-square test*
- goodness of fit

2. parametric tests

- *t-test*
- *ANOVA (analysis of variance)*
- *regression and correlation*
- time series
- index numbers

Parametric and non-parametric tests are available for testing hypotheses related to differences in the metric and non-metric scales. In the parametric case, the *t-test* is used to examine hypotheses related to the population means.

Thank you

Your ? Through email