CHAPTER 1

BASIC ISSUES CRUCIAL IN SETTING RESEARCH DESIGN

Basic Steps to Follow

Researchers usually follow similar research steps in conducting their research. The research activities can be categorized under the following headings:

- **Focusing on a problem:** The first step is to identify a general problem area. The topic should be of interest to the researcher or should be related to his/her own field of expertise. There may be a problem for which the researcher is in *need of finding a solution*. Problems can be derived from theory or from one's personal experiences (see Vol. 1 Chapter 2).

- **Formulating a purpose statement or research question:** The next step is to narrow down the general problem so that it would be specific enough to conduct the research on. After a library search and initial reading, one starts narrowing down the topic. This can be done by clearly formulating a specific question regarding the problem. A research problem is considered acceptable if investigation can be done within the given time, on the base of collection and analysis of data which can be obtained from available sources. In other words, by formulating a research question or giving a specific statement focused on a specific problem, the researcher finds himself/herself in a position of stating the purpose and the scope of the research (see Vol. 1 Chapter 3).

- **Giving the background and the rationale of the research:** At this stage, general background information and the theory the research is going to be based on is determined. Generally, in this section, the researcher gives references crucial to the research. This section acts as a transition between the research to be conducted and the prior research. In other words, the researcher, by giving a summary of the rationale underlying the research to be conducted, tries to provide a *justification* for the investigation of the proposed problem.

- **Formulating a hypothesis:** In starting out a research project, one has to have some *expectations* regarding the results of the research. Each tentative explanation of these expectations is defined to be the statement of a hypothesis (see Vol. 1 Chapter 3).

- **Reviewing the literature:** Once the research question is formulated, the literature in that field is reviewed in depth. This is especially true in writing a thesis or a dissertation where a thorough survey of the related literature is needed. Thus, the researcher has to give detailed information on the *theoretical base of the study.*
  If the answer is found in the literature, the research design does not go beyond a survey. If the answer still remains vague, a research method to investigate the problem has to be looked into. While reviewing the literature, one has to be aware of note-taking procedures and storing bibliography cards (see Vol. 1 Chapter 4).

- **Choosing the research design:** After formulating the problem and the hypothesis, the researcher chooses a research design that will best suit the purpose of the study. Now is

the time for the researcher to decide on the *techniques and instruments to be utilized for obtaining information* related to the problem and the hypotheses (see Vol. 1 Chapter 5).

- **Making a detailed plan of the study:** When the research design to be applied becomes definite, a detailed plan of the study has to be made. This is first done in the form of an outline (see Vol. 1 Chapter 2). Once the rough *outline* is made, the researcher can work towards the analysis, trying to consider the scope, the hypotheses, the specific objectives of the study, and the method of data collection and analysis.

  As an indication of this overall plan, graduate students are required to submit a research proposal as a partial fulfillment of a degree they are seeking (see Vol. 1, Chapter 6). After having reviewed the related literature, the students write their *research proposal* focusing on the statement of the problem, the background of the study, the hypotheses made, and the research design to be implemented.

  All the steps mentioned above enable the researcher to obtain enough knowledge to design the research to be conducted. At university, the layout of the design is written for approval in the form of a research proposal. Once the proposal is approved by the research committee, the following steps need to be taken into consideration in order to complete the research, which will be the concern of this volume:

- **Collecting data:** Once the research design gets its final form, the researcher starts collecting data or information that will shed light on the problem. The method of data collection varies depending on the problem and the design of the research (see Chapter 2).

- **Analyzing the data:** The collected data has to be categorized and analyzed in parallel to the research design to be adopted (see Chapter 3).

- **Generating conclusions:** This is the last step of the research. At this stage, the researcher tries to arrive at a conclusion on the basis of the analysis of the collected data. The conclusion also includes a discussion on the fulfillment or rejection of the hypothesis formulated at the early stages of the research. Reasons for the fulfillment or the rejection of the hypothesis should be explicitly stated (see Chapter 3).

- **Organizing the collected information:** The collected information is organized depending on the context. In order to do a successful job, the researcher has to be aware of the organizational patterns in writing and the means of combining ideas together to compose a coherent and cohesive piece of writing (see Chapter 4). During the organization of the information, the research outline is revised according to the notes obtained from the literature review.

- **Writing a thesis, a dissertation, or a research article:** This is where one needs professionalism and expertise. For that reason, one has to be aware of some commonly used phrases and structures utilized in formulating the information at different sections of the research (see Chapter 5). After the first draft, the research study is written in its full form, then revised and edited thoroughly (see Chapters 6 and 7).

  The writing style to be applied is very important as well. There are three main formats adopted by scholars: APA (see Appendix A), MLA (see Appendix B), and Chicago Manual Style. The form commonly used in social sciences is the APA format, and this book applies the same format as well. When writing an article for a journal, the researcher must adopt the format required by that particular journal. Most of the journals indicate the required format generally inside the cover page.

Basic Concepts to Consider

In every stage of the research (formulating the problem, the purpose statement, or the hypothesis; in describing the data in graphic representation or analyzing them in statistical terms) the researcher has to have a broad perspective about the population and the samples to be chosen, as well as the variables in order to be able to determine the effect of the independent variable(s) on the dependent variable(s) (see Vol. 1, Chapter 3 for detailed information).

Population

When choosing the population the researcher should ensure that the individuals included in the population have the same characteristics except the one(s) stated in the variables. In other words, the other variables should be kept constant in order to arrive at a sound conclusion. For instance, the population may be kept the same in reference to age, sex, and educational background depending on the purpose of the research.

The population may be any *size* and may cover any *attribute* and any *geographical* area. What is important, however, is that the generalizations made at the end of the study should be limited to the chosen population. For instance, the results of a study conducted at one school in Ankara cannot be generalized to the school system in that city. For that reason, "the key is to define your population in sufficient detail so that others may determine how applicable your finding might be to other situations" (Gay, 1987, p. 103).

In order to make a generalization about the whole population, samples representing different sections of the population need to be included. In most cases, the population to derive generalizations from is hardly available. Therefore, the researcher realistically selects the population from what is available.

For instance, if the interest is to find out about the writing skills of children at an elementary school, then the population should comprise elementary school children. This characteristic (elementary school) is too broad for the study to be conducted; therefore, the population should be narrowed down by adding extra characteristics that would limit the scope of the study (see also Chapter 2). For this purpose, another type of limitation such as focusing only on one grade in elementary schools can be set. This might still be broad. Then, the population can be narrowed down by considering all the sixth-graders (a) in the country, (a) in a particular geographical region, (c) in a city, (d) in a district in the city, (e) in two elementary schools. This indicates that when the *target population* (see Example 1) cannot be reached at, attempt is made to collect data from the *accessible population* (see Example 2). When the accessible population is found to be too broad; then, the researcher chooses a *sample* (see Example 3) among the accessible population.

Examples:

1) Target population:	All the sixth-graders in Turkey.
2) Accessible population:	All the sixth-graders in Adana.
3) Sample:	300 fifth-graders selected from different elementary schools in Adana.

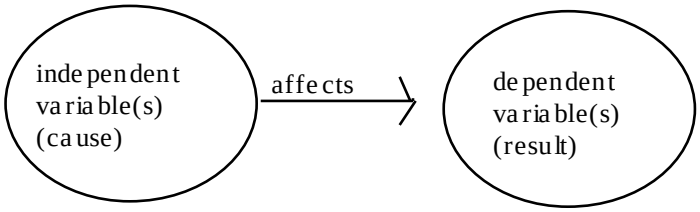
Sampling

After having decided how many subjects should be included in the sample, the researcher should make an effort to select these subjects so that they would accurately represent the population characteristics. There are different methods of random sampling. The most commonly cited ones are simple, stratified, cluster, and systematic random sampling (see Vol. 1, Chapter 3 for details).

Basic Variables

The independent variable is the major variable which is to be investigated. It is the variable which is selected, manipulated, and measured by the researcher. It is the independent variable which is put into action to serve as a stimulus within the treatment. For that very reason, it is also referred to as input, manipulated, treatment, and stimulus variable.

The dependent variable is the variable which is observed and measured to determine the effect of the independent variable. It is called dependent variable because the outcome of the research depends on how the independent variables are managed or manipulated.



There are two kinds of descriptive statements regarding variables:
1) a state description
2) a process description

A state description specifies a value on a variable. For instance, Value b is the observed value on variable B. A process description, on the other hand, specifies the causal relationship between two variables (see also Vol. 1, Chapter 5). For instance, when different values on variable A are produced , different values on variable B are observed. The independent variable requires process definition because the experimenter selects the variable and manipulates it to determine the effect of this variable on other variables that are kept constant. The dependent variable has a measurement type of operational definition given in scores or numerical values and thus requires state description.

A variable is a set of mutually exclusive properties. In other words, a variable is an attribute of a person or an object which varies from person to person or from object to object. The variable can be coded alphabetically, numerically, or in an alphanumeric style (combination of the first two). For instance, if there are two levels of the variable (1= sex, 2= smoker, nonsmoker), each level could be coded as M= 0, F=1. Symbols such as figures, letters, or words are referred to as numerals rather than numbers. Social security numbers, telephone numbers, and serial numbers on library cards are considered numbers because the digits do not indicate quality. They are used for coding purposes. The numerals indicating the variables serve for the same purpose.

Extraneous Variables and Their Threat

to Different Experimental Designs

Lack of information on the characteristics and function of variables could lead to questionable results because if these variables are not taken into consideration properly, the internal and/or external validity of the research is threatened.

By means of the independent variable, the internal validity of the study is controlled. In other words, the internal validity is related to the outcome of the function of the factor selected rather than other uncontrolled factors. The researcher, however, has to control other background variables to cancel out, reduce, or minimize the effects of these extraneous variables on the outcome of the study. There are eight different classes of extraneous variable, which, if not controlled, might produce effects confounded with the effect of the factor selected. Therefore, these uncontrolled extraneous variables may become threats to the study (see also Vol. 1, pp. 90-93). Following are some threats and their descriptions:

History	The specific events occurring between the first and second measurement in addition to the independent variable selected for the experiment.
Maturation	The mental and physical changes that occur within subjects.
Testing	The effects of taking a test upon the scores of a second testing.
Instrument Decay ...	Changes in the calibration of a measuring instrument or changes in the observers or scorers involved in the experiment.
Subject Character ...	Braces resulting in differential selection of respondents for the comparison groups
Mortality	Differential loss of respondents from the comparison groups.
Implementation	As a result of unplanned treatment, selection-maturation interaction might be mistaken for the effect of the independent variable.
Location	Reactive effects of experimental arrangements which would preclude generalization about the effect of the experimental variable upon persons being exposed to it in nonexperimental settings.

Reviewing Different Experimental Designs in Statistical Terms

Experimental designs take different names depending on the way the experiment is conducted on the sample population:

1. The One-shot Case Study

X O

X refers to the treatment, in other words, the exposure of a group to an experimental variable or event, the effects of which are to be measured.

O refers to some process of observation or measurement.

These symbols are used to indicate the process that takes place in each type of experimental designs. While providing the illustrations, the degree of control of extraneous

variables that are involved in each design will be mentioned (see also Vol. 1, Chapter 5, Part II for details).

The left-to-right dimension of these symbols indicates temporal order, and each occurrence is indicated with a different number. Simultaneous occurrences of **X**s and **O**s are indicated vertically.

2. The One-group Pretest-posttest Design

$$O_1 \quad X \quad O_2$$

By the above formula, the reader is informed that first there has been a process of observation or measurement, then a treatment has been given. And after the treatment a second process of observation or measurement has taken place.

In this type of study, there is no control; therefore, the study has no scientific value. This design may not have control on the following threats:

History: Between O_1 and O_2 many other change-producing events may have occurred in addition to X.

Maturation: All biological or psychological processes systematically vary with the passage of time, independent of specific external events.

Testing: The effect of the pretest itself is the extraneous variable. On achievement and intelligence tests, students taking the test for a second time usually do better than those taking the test for the first time.

Instrument Decay: Autonomous changes in the measuring instrument might account for an O_1 - O_2 difference. In cases when human observers are used, learning, fatigue within the observers, may produce the O_1 and O_2 difference. The grading standards may shift between O_1 and O_2 .

3. The Static-group Comparison

$$\begin{array}{ccc} X & & O_1 \\ \text{—} & \text{—} & \text{—} \\ & & O_2 \end{array}$$

Subject: A group which has experienced X is compared with another group that has not gone through a similar experience just for the purpose of establishing the effect of X. There is no means of certifying that the groups would have been equivalent had it not been for X. A dashed line separating two groups means no selection has been made. If O_1 and O_2 differ, this difference could have come about through the differential recruitment of subjects making up the groups. In other words, the groups might well have differed anyway, without the occurrence of X.

Mortality: The production of O_1 - O_2 differences in groups may be due to the differential drop-outs of persons from the groups.

4. The Pretest-posttest Control Group Design

R	O_1	X	O_2
R	O_3		O_4

The symbol **R** refers to random assignment. Equivalent groups are formed by randomization but still there are some variables that may not be controlled.

History: General historical events that might have produced an $O_1 - O_2$ difference would also produce an $O_3 - O_4$ difference.

Maturation and testing: They should be manifested equally in experimental and control groups.

Regression: Regression is controlled (as far as mean differences are concerned) no matter how extreme the group is on pretest scores. If both experimental and control groups are randomly assigned from this extreme pool, the control group regresses as much as does the experimental group.

5. The Solomon Four-group Design

R	O_1	X	O_2
R	O_3		O_4
R		X	O_5
R			O_6

In this new design, aside from the elements in Design 4 (O_1 through O_4), two more groups (experimental and control groups lacking the pretest) are added into the study. The first and the third groups form the experimental groups, and therefore receive some treatment. The second and the fourth groups, on the other hand, form the control groups. While all groups are posttested, only the first and the third groups are pretested. With this new design, the researcher can determine the effects of testing as well as the interaction of testing. The effect of X replicated in four different fashions: $O_2 > O_1$, $O_2 > O_4$, $O_5 > O_6$, $O_5 > O_3$.

6. The Posttest-only Control Group Design

R	X	O_1
R		O_2

A pretest is not given to any of the groups. There is control of the subject choice due to randomization.

7. The time-series experiment

O_1	O_2	O_3	O_4	X	O_5	O_6	O_7	O_8

In this design, there is a periodic measurement process on some group of individuals and an introduction of an experimental change into this time series of measurements.

8. The Equivalent Time Series Design

X_1 O X_0 O X_1 O X_0 O

A recurrent form of one-group experimentation employs two equivalent samples of occasions in one of which the experimental variable is present while in the other it is absent.

History : The major weakness (something could happen between the pretest and the posttest) is controlled by presenting X on numerous separate occasions by rendering any unlikely rival explanation based on the coincidence of these extraneous events.

9. The Equivalent Materials Design

M_aX_1O M_bX_0 O M_cX_1O M_d X_0O

In this design, there is an application of equivalence of samples of materials to which the experimental variables are compared. The sample Ma and Mc, in sampling terms, are equal to sample Mb and Md.D.

10. The Nonequivalent Control Group Design

O X O

O O

An experimental group and a control group are both given a pretest and a posttest without having pre-experimental sampling equivalence. Instead, the groups constitute naturally assembled collectives such as classrooms, as similar as availability permits but yet not so similar that one can dispense with the pretest. The assignment of X to one group or the other is assumed to be random and under experimental control.

This design should not be confused with Design 4 (the Pretest-posttest Control Group Design), in which experiment subjects are assigned randomly from a common population to the experimental and control groups. If experimental and control groups are similar, this design controls the main effects of history, maturation, testing, and instrumentation.

11. The Counterbalanced Design

Groups	Time 1	Time 2	Time 3	Time 4
Group A	X_1O	X_2O	X_3O	X_4O
Group B	X_2O	X_4O	X_1O	X_3O

Group C	X_3O	X_1O	X_4O	X_2O
Group D	X_4O	X_3O	X_2O	X_1O

X= experimental treatment O= posttest

The control in the experiment is achieved by entering all respondents into all treatments in different orders. Four experimental treatments are applied in a randomized manner in turn to form four naturally assembled groups. The design has been diagrammed with posttests only.

The design contains three classifications: groups, occasions (different times), and experimental treatments. Each classification is "orthogonal" to the other two, in which each variant of each classification occurs equally often with each of the variants of the other classifications. Each experimental treatment, in statistical terms, each X occurs only once in each column and only once in each row; therefore, the order of these treatments to different groups varies.

The researcher determines the effectiveness of the various treatments "simply by comparing the average scores for all groups on the posttest for each treatment. In other words, the average of the posttest scores for all groups for Treatment 1 (X_1) can be compared with the average of the posttest scores for all groups for Treatment 2 (X_2), and this procedure is applied for all the treatments within the study (Fraenkel & Wallen, 1990, p. 243). This design controls the subject characteristics threat, but it is affected by the multiple-treatment interference. In other words, students can be affected by the previous treatments while performing on the following treatments.

12. The Separate-sample Pretest-posttest Design

R	O	(X)	
R		X	O

The rows represent randomly equivalent subgroups. The parenthetical X stands for a presentation of X irrelevant to the argument. One sample is measured prior to the X, an equivalent one subsequent to X. In other words, there is only one treatment for two groups. One group is tested before the treatment, the second is tested after the treatment.

13. The Separate-sample Pretest-posttest Control Group Design

R	O	(X)	
R		X	O
---	---	---	---
R	O		
R			O

This is almost the same as Design 12. Only a control group is added.

14. The Multiple Time-series Design

O_1	O_2	O_3	O_4	X	O_5	O_6	O_7	O_8

O_1 O_2 O_3 O_4 O_5 O_6 O_7 O_8

This design is similar to Design 7. There is an addition of another set of groups with no introduction of treatment. This is to observe the effects of the treatment under an unbiased condition.

If we have good background information about the basic steps and concepts related to research, we can easily set a proper research design that would enable us to collect and organize our data. Moreover, we would feel more confident in analyzing and interpreting the organized data in statistical terms.

EXERCISES

A. The administrators of television station ATV wanted to find out the two most popular programs televised through their channel. For that purpose they interviewed a random group of 1000 TV owners in Ankara.
   1. What is the population?
   2. What is the sample?
   3. What type of random sampling technique did they use?

B. How would you collect data to conduct the following studies?
   1. A study of the effect of language laboratory in developing students listening skills in second language learning.
   2, A study of the effect of knowing the use of computer in getting jobs.

C. Find a newspaper or a magazine that uses statistics. Write down the population and/or the sample taken into consideration for that particular study.

D. Suppose you are going to make a statistical profile of your class. Which qualification would you take into account? Try to classify these qualifications as physical properties, course load, parental information, residential information, etc.

E. Suppose you are the owner of a private bus company, and you want to make a survey before you make the bus schedule. What would you inquire first? How would you conduct this study?

F. Suppose you opened a pastry shop and started selling different types of desserts? What would you base the daily production of these desserts on?

G. Suppose you are conducting a survey to find out the opinions of people on a specific issue on education. Would the following be a good technique of collecting data? If your answer is no, state your reason.
   1. Select every fourth woman entering a beauty shop.
   2. Select every fourth man entering a football match.
   3. Select every fourth person entering a theater.

H. Suppose you are the owner of a big supermarket, and a truck containing coffee cups and other ceramic plates has arrived. How would you get a random sample of 10 boxes to check if any has been broken?