



# Methods and Techniques of Social Research

A.J.M. Sufian<sup>1</sup>

<sup>1</sup> Department of Management & Marketing, College of Business, University of Bahrain, Kingdom of Bahrain

Received 20 March 2015, Revised 28 April 2015, Accepted 29 April 2015, Published 1<sup>st</sup> July 2015

**Abstract:** The purpose of methods and techniques of social research is to offer tools necessary for dealing with social problems through research in a scientific manner. The special emphasis of methods and techniques of social research arises mainly in the coverage of its materials. It contains both methods of collecting observations pertaining to social phenomena and techniques for analyzing them. The greatest advantage of having both methods and techniques within the same covers is that a researcher may easily establish a direct linkage between them and this may help greatly to choose an appropriate technique to analyze data.

**Keywords:** ANOVA, Data processing, Regression, Test of Hypothesis.

## 1. INTRODUCTION

The book [18] is organized into three parts. Part 1 deals with empirical research and its structure. Part 2 is concerned with different methods of collecting data while part 3 deals with data processing and some important techniques of data analysis.

## 2. EMPIRICAL RESEARCH

In practice, the accumulation of evidence for or against any particular theory involves planned research designs for the collection of empirical data. Several typographies for such designs have been suggested, one of the most popular of which comes from [1], [2], [3], [4] and [5]. They are responsible for popularizing the widely cited distinction among pre-experimental, experimental, and quasi-experimental designs and are staunch advocates of the central role of randomized experiments in educational research. In particular, they view the experiment: as the only means for settling disputes regarding education practice, as the only way of verifying educational improvements, and as the only way of establishing a cumulative tradition in which improvements can be introduced without the danger of a faddish discard of old wisdom in favor of inferior novelties.

The hallmarks of an experiment to Campbell and Stanley, among others are (a) random assignment of cases to comparison groups, (b) control of the implementation of a manipulated treatment variable, and (c) measurement of the outcome with relevant, reliable

instruments. Controlled experimentation allows for replication of the conditions of the experiment so that independent researchers can attempt to repeat the results of the experiment. In contrast, non-experimental studies may use convenience samples, comparison groups formed by post-hoc matching and similar procedures. [3] provide a framework for evaluating the limitations that various types of research studies pose with respect to inferring a causal link between independent (treatment) and dependent (outcome) variables. They posit a necessary relationship between the validity of an individual research study and the generalization of results from this study to wider populations. They argue that: Internal validity is the basic minimum without which any experiment is uninterruptable. Did in fact the experimental treatments make a difference in this specific experimental instance?.

Typical of potential threats to internal validity are: (1) uncontrolled, extraneous events occurring during the study (called a "history" threat); (2) failure to randomize interviewers or raters across comparison groups (called an "instrumentation" threat); (3) biased or differential selection of cases as occurs when groups are self-selected in a case-control study (call a "selection" threat); and (4) differential loss of cases from comparison groups when there is no pretest to assess the impact of the loss (called an "experimental mortality" threat). Additional threats are discussed in [3]. Note that, in general, control of threats to internal validity allows the research to rule out plausible rival hypotheses concerning differences between comparison groups.



**Randomization:** Ideally, subjects should be randomly selected from the target population and then randomly assigned to treatment conditions. Internal validity (though not external validity) can be attained if available samples are randomly assigned to treatment conditions.

Quasi-experimental designs such as cohort studies require pre-measures and other covariates that allow for statistical adjustment in an attempt to control for history and other threats. Similarly, case-control studies require covariates for adjustment purposes.

However, it should be noted that adjustment for all relevant, non-randomized competing causes in non-experimental studies is an essentially hopeless task. [10] notes that analysis of covariance (ANCOVA) can be used for increasing precision in experimental studies and for attempting to adjust for initial differences in non-experimental studies. The application of ANCOVA for the first purpose is well founded, and may prove useful in diverse research areas. The applications of ANCOVA for the second purpose, however, is highly questionable because it is fraught with serious flaws. Unfortunately, application of ANCOVA in quasi-experimental and nonexperimental research is by and large not valid.

**Control:** Extraneous factors associated with variation in an outcome variable can be controlled by techniques such as selection, stratification, and possibly statistical adjustment or can be randomized. For example, if there are known socio-economic status (SES) differences on a dependent variable, the researcher can: (a) select cases within a relatively narrow range of SES so that its impact becomes negligible or, at least, lessened; (b) stratify experimental cases into SES blocks that can be incorporated into the design and analysis; or (c) obtain a suitable measure of SES and partial out its influence. In experimental settings, the benefit of all of these procedures is to reduce unexplained within-group variation and, thereby, both increase the likelihood of detecting an effect (i.e., increase power) and reduce the uncertainty associated with the magnitude of an effect (i.e., decrease the width of confidence intervals). Alternatively, the research can ignore SES differences, randomly assign cases to groups, and lose the above benefits.

**Reliability:** It is preferred that outcomes (and covariates) be assessed with relatively little measurement error. Other things being equal, unreliability increases unexplained variation within groups and reduces the power of the analysis. In practice, it may be impractical to assess the reliability of measurement procedures within the scope of a given study, but the selection of measurement instruments should certainly take this factor

into consideration. On the other hand, if a study involves observations or ratings by judges, some effort must be undertaken to assure consistency of measurement across judges; see, [6], [7], [8], [9], [10] and [11].

**Validity:** In selecting a relevant measure for an outcome variable, it is critical that logical inferences can be made from the operationalizations upon which the measure was based to the theoretical constructs relevant to the study. Construct validity refers to the degree to which inferences of this type can legitimately be made; see, [12], [13], [14], [15].

**Implementation of Treatment Variable:** An overlooked consideration in many studies is the provision of evidence that the independent variable of interest has actually been applied as intended. [17] described a famous failure of implementation. In 1930 in Scotland the Department of Health conducted the Lanarkshire Milk Experiment to investigate the advantage of giving extra milk to schoolchildren. The experiment, involving 20,000 children, was seriously compromised by some teachers who gave the extra milk to students they considered most needy as opposed to those selected by randomization. The lesson is that there must be some record or documentation supporting the fact that the intended treatment has taken place.

**Analysis Issues:** Research studies without serious design limitations may nevertheless suffer from inadequate or inappropriate analyses. While there are often alternative analytical approaches that result in equivalent analyses with respect to interpretation of results, it is also the case that inappropriate analysis may limit interpretability. Among issues that arise reasonably often are: (a) failure to utilize an appropriate unit of analysis (e.g., ignoring nesting of students within schools and employing ordinary ANOVA when hierarchical linear modeling would be more appropriate); (b) arriving at models by exploratory procedures but interpreting results as if models were confirmed (e.g., using stepwise multiple regression to "confirm" the importance of predictor variables or using model modification indices in structural equation modeling to alter an initial model to improve fit to data); (c) deriving estimates from complex survey designs without considering design issues (e.g., neither using weighted estimates nor modeling the design when analyzing NAEP data); and (d) ignoring distributional assumptions with parametric procedures such as multiple regression, ANOVA, structural equation modeling, etc. (e.g., ignoring the impact of outliers, extremely skewed distributions of residuals, or lack of homogeneity of variance). There are, of course, many more subtle issues such as the mistaken notion that non-

parametric tests for location (e.g., Mann-Whitney U) are insensitive to homogeneity of variance assumptions.

**Interpretation Issues:** While the use of inferential statistical methods has been a valuable tool in many applied research fields, their use has also led to some unfortunate opportunities to make incorrect or misleading interpretations of results. Recent emphasis on reporting effect sizes may be viewed as valuable, but all too often this takes the form of comparing a computed effect size (e.g., standardized absolute mean difference) with some completely arbitrary standard (e.g., .5 as indicating a "medium" effect). In fact, a statistically significant outcome for, say, a two-independent-sample t test for means merely suggests that the result is "surprising" when compared to a model of chance variation. The practical interpretation of the observed outcome must be made within the context of the research setting.

Where Part 1 includes the first four chapters. Chapter 1 discusses basic issues and then briefly examines the different interrelated steps in social research. Chapter 2 is devoted to the development of research problem and relevant hypotheses. A test of such hypotheses usually provides solution(s) to the problem that initiated the research. Once the hypotheses are developed, measurements of variables embedded in those hypotheses as well as of other relevant variables are needed. Chapter 3 deals with such measurements. This chapter includes development of indexes and scales for measuring concepts that lack empirical referents. Validity and reliability which are inherently linked with measurements are taken up in this chapter. Once the measurements of variables are developed, a group of people known as respondents are needed to collect data on these variables from. Chapter 4 deals with the methods of scientific sampling to select these respondents; see, [16], [17] and [18].

### 3. DIFFERENT METHODS OF COLLECTING DATA

Data Collection is an important aspect of any type of research study. Inaccurate data collection can impact the results of a study and ultimately lead to invalid results. Data collection methods for impact evaluation vary along a continuum. At the one end of this continuum are quantitative methods and at the other end of the continuum are Qualitative methods for data collection .

The **Quantitative data collection methods** rely on random sampling and structured data collection instruments that fit diverse experiences into predetermined response categories. They produce results that are easy to summarize, compare, and generalize. Quantitative research is concerned with testing hypotheses derived from theory and/or being able to estimate the size of a phenomenon of interest.

Depending on the research question, participants may be randomly assigned to different treatments. If this is not feasible, the researcher may collect data on participant and situational characteristics in order to statistically control for their influence on the dependent, or outcome, variable. If the intent is to generalize from the research participants to a larger population, the researcher will employ probability sampling to select participants.

Typical quantitative data gathering strategies include:

- Experiments/clinical trials.
- Observing and recording well-defined events (e.g., counting the number of patients waiting in emergency at specified times of the day).
- Obtaining relevant data from management information systems.
- Administering surveys with closed-ended questions (e.g., face-to face and telephone interviews, questionnaires etc).

**Interview:** In Quantitative research (survey research) ,interviews are more structured than in Qualitative research. In a structured interview,the researcher asks a standard set of questions and nothing more.(Leedy and Ormrod, 2001)

**Face -to -face interviews** have a distinct advantage of enabling the researcher to establish rapport with potential participants and therefore gain their cooperation. These interviews yield highest response rates in survey research. They also allow the researcher to clarify ambiguous answers and when appropriate, seek follow-up information. Disadvantages include impractical when large samples are involved time consuming and expensive; see, [7], [20] and [21].

**Telephone interviews** are less time consuming and less expensive and the researcher has ready access to anyone on the planet who has a telephone. Disadvantages are that the response rate is not as high as the face-to-face interview as but considerably higher than the mailed questionnaire. The sample may be biased to the extent that people without phones are part of the population about whom the researcher wants to draw inferences.

**Computer Assisted Personal Interviewing (CAPI):** is a form of personal interviewing, but instead of completing a questionnaire, the interviewer brings along a laptop or hand-held computer to enter the information directly into the database. This method saves time involved in processing the data, as well as saving the interviewer from carrying around hundreds of questionnaires. However, this type of data collection



method can be expensive to set up and requires that interviewers have computer and typing skills.

#### 1) Questionnaires

**Paper-pencil-questionnaires** can be sent to a large number of people and saves the researcher time and money. People are more truthful while responding to the questionnaires regarding controversial issues in particular due to the fact that their responses are anonymous. But they also have drawbacks. Majority of the people who receive questionnaires don't return them and those who do might not be representative of the originally selected sample. (Leedy and Ormrod, 2001)

**Web based questionnaires** : A new and inevitably growing methodology is the use of Internet based research. This would mean receiving an e-mail on which you would click on an address that would take you to a secure web-site to fill in a questionnaire. This type of research is often quicker and less detailed. Some disadvantages of this method include the exclusion of people who do not have a computer or are unable to access a computer. Also the validity of such surveys are in question as people might be in a hurry to complete it and so might not give accurate responses. (<http://www.statcan.ca/english/edu/power/ch2/methods/methods.htm>)

Questionnaires often make use of Checklist and rating scales. These devices help simplify and quantify people's behaviours and attitudes. A **checklist** is a list of behaviours, characteristics, or other entities that the researcher is looking for. Either the researcher or survey participant simply checks whether each item on the list is observed, present or true or vice versa. A **rating scale** is more useful when a behaviour needs to be evaluated on a continuum. They are also known as Likert scales. (Leedy and Ormrod, 2001).

**Qualitative data collection methods** play an important role in impact evaluation by providing information useful to understand the processes behind observed results and assess changes in people's perceptions of their well-being. Furthermore qualitative methods can be used to improve the quality of survey-based quantitative evaluations by helping generate evaluation hypothesis; strengthening the design of survey questionnaires and expanding or clarifying quantitative evaluation findings. These methods are characterized by the following attributes:

- they tend to be open-ended and have less structured protocols (i.e., researchers may change the data collection strategy by adding, refining, or dropping techniques or informants)

- they rely more heavily on interactive interviews; respondents may be interviewed several times to follow up on a particular issue, clarify concepts or check the reliability of data
  - they use triangulation to increase the credibility of their findings (i.e., researchers rely on multiple data collection methods to check the authenticity of their results)
  - generally their findings are not generalizable to any specific population, rather each case study produces a single piece of evidence that can be used to seek general patterns among different studies of the same issue
- Regardless of the kinds of data involved, data collection in a qualitative study takes a great deal of time. The researcher needs to record any potentially useful data thoroughly, accurately, and systematically, using field notes, sketches, audiotapes, photographs and other suitable means. The data collection methods must observe the ethical principles of research.

The qualitative methods most commonly used in evaluation can be classified in three broad categories:

- **indepth interview**
- **observation methods**
- **document review**

Different ways of collecting evaluation data are useful for different purposes, and each has advantages and disadvantages. Various factors will influence your choice of a data collection method: the questions you want to investigate, resources available to you, your timeline, and more.

Where A variety of methods for collection of data depending on the nature of a given research, hypotheses to be tested, and resources available have been developed and are discussed in part 2 that comprises of chapters 5, 6, 7, and 8. Chapter 5 deals with the experimental method that involves performing a certain act and then observing the change on a given variable as a consequence of that act. Chapter 6 presents the field method in which information about social phenomena are collected in their natural settings. Chapter 7 is concerned with content analysis and use of data from secondary sources. That is, it deals with materials that already exist. Chapter 8 deals with the survey research method. It mainly discusses issues relevant to the construction of questionnaires and different types of interviews. The advantages and disadvantages of each method of collecting information have been discussed in these chapters with a view to familiarizing the reader with the strengths and weaknesses of these methods. This may help him choose an appropriate method in a given research situation.



#### 4. DATA PROCESSING AND TECHNIQUES OF DATA ANALYSIS

The processing of data/information is an essential dimension of stream lining the facts and writing of a field report. A separate account of processing is given here.

##### Processing of primary data:

The primary data collected from the field remains in the raw form of statements, digits and qualitative terms. The raw data contains error, omissions and inconsistencies. It requires corrections after careful scrutinizing the completed questionnaires. The following steps are involved in the processing of primary data.

##### Editing of data:

The editing of data can be done at two stages: field and post-field editing. The field editing is a review of reporting by the investigator for completing what has been written in an abbreviated form during interviewing the respondent. The post-field editing is carried out when field survey is completed and all the forms of schedule have been collected together. This type of editing requires review of all forms thoroughly.

##### The coding of data:

To keep the response with in limited alternatives, we need to assign some alphabetical or numerical symbols or both to the answers. The alternatives must be mutually exclusive i.e. defined in one concept or term only. This form of processing is known as coding. For example in a question of educational qualifications alternative choices given are: Uneducated; Below Matriculation; Matriculation & above but below Graduate; Graduate & above; Technical Diploma; Technical Degree.

The alphabetical codes assigned to these alternatives could be A,B,C,D,E and F. Similarly, numerical codes to these alternatives could be 1,2,3,4.,and 5 respectively. It is necessary for the efficient analysis.

Though coding exercise is a part of the formulation of questionnaire yet responses to questions need to be coded and made final at the processing stage. This simplifies the transfer of data from questionnaires to the master chart. It is a two dimensional chart in which observations are entered on one axis (X) and details of the responses on the other axis (Y). The calculations becomes easier and quicker if the details are coded and entered in the master chart or fed in the computers.

**Organization of Data:** The data information collected through different sources should be organized. The first task in this regard is to develop a master chart. For example in a local area survey, we record individual

households in rows and the details of population, function, facilities and amenities etc. in columns. Thus a large chart is prepared that contains, practically, all relevant information/data. Finally the total of rows and columns are cross-checked. The information arranged in an ascending order is known as the array of data. The set of information related to specific entity is called the field. The following illustration demonstrates the way data is organized.

**Classification of data:** A huge volume of raw data collected through field survey needs to be grouped for similar details of individual responses. The process of organizing data into groups and classes on the basis of certain characteristics is known as the classification of data. Classification helps in making comparisons among the categories of observations. It can be either according to numerical characteristics or according to attributes.

**Presentation of data:** The presentation of data could be tabular, statistical and cartographic. In case of tabular form of presentation, data related to different variables should be classified and compared. Various statistical techniques are available to derive accurate and precise results. Since techniques have a large range coupled with the limitations of their own, selection of appropriate technique needs to be made for the purpose. The construction of graphs, charts, diagrams and maps are the various forms of cartographic presentations. The data is transformed into cartographic system which is used for visual presentation.

**Tabular Presentation:** It is used for summarization of data in its micro form. It helps in the analysis of trends, relationship and other characteristics of a given data. Simple tabulation is used to answer question related to one characteristic of the data whereas complex tabulation is used to present several interrelated characteristics. Complex tabulation results in two way, three way tables which give information about two or three inter-related characteristic of data.

##### Some statistical analysis is mainly applied to:

- Sampling techniques
- Regression Analysis
- Factor Analysis
- Correspondence Analysis
- Multidimensional Scaling

The next task of the researcher is to process and analyze the data - subject matter of part 3. A number of techniques of data analysis depending on the structure of variables are available. These techniques may range from simple univariate description to complex multivariate



procedures. Attempt has been made to present each technique in a way that will help the reader develop the skill necessary to know when and how to use it. Chapter 9 is concerned with data processing - an act necessary for preparing the data for analysis. Included here are such topics as editing, coding, and construction of code categories. Chapter 10 introduces univariate analysis that involves creating frequency distribution and developing summary measures. The normal curve is also discussed in this chapter. Bivariate and multivariate analysis for both nominal and ordinal variables are presented in chapters 11 and 12 respectively. Chapter 11 deals with measures of relationships among such variables while chapter 12 shows how to control the effects of extraneous variables on these relationships through elaboration approach. When two variables are found to be related, the relationship may be true, spurious, or an obscured one. It is essential that before the researcher concludes that the relationship is true, he makes further investigation to rule out the possibility that it is spurious or obscured. Spuriousness refers to a fake relationship between two variables that arises when they are not actually related but the presence of one or more extraneous variables produces such a relationship between them. If those extraneous variables are somehow not allowed to affect the two variables under consideration, the fake relationship disappears. Similarly, these extraneous variables may obscure the relationship and once again, if they are not allowed to affect these two variables, the original relationship surfaces. The process by which the extraneous variables are not allowed to affect the two variables is known as 'controlling their effects'. In experimental designs, this control is achieved by matching or by randomization while selecting the respondents, and then allocating them into two groups - experimental and control. There are a number of methods for controlling the effects of extraneous variables through statistical adjustments. The most basic of these methods is known as elaboration that involves comparison of subgroups. This method achieves control by the use of contingency table analysis when the variables are measured at nominal level. An understanding of this method will help the reader grasp the basic logic of the multivariate analysis in a very straightforward manner. Chapters 13 and 14 present respectively bivariate and multiple regression techniques - techniques which are used when all the variables are measured at interval level. These chapters present mainly the regression model, estimation of parameters and their interpretations, measures of relationships between variables, and tests of significance. Analysis of residuals which is relevant to both bivariate and multiple regressions is presented only in chapter 13 since the basic concepts are the same in both cases. The addition of more than one independent variable in multiple regression entails some new topics such as interaction effects, multicollinearity,

dummy variable regression, and are presented in chapter 14. Chapter 15 presents the logistic regression technique. This technique is used when the dependent variable is a nominal dichotomous variable. The topics mainly discussed here are prospective and retrospective studies, odds and odds ratio, estimation of parameters, and likelihood ratio test. Chapter 16 deals with analysis of variance and covariance. Finally, discriminant analysis is discussed in chapter 17. It is a technique of studying the relationship between a nominal variable and a set of interval variables. The categories or groups of the nominal variable must be mutually exclusive in that, any individual can belong to one of the groups only. We are basically interested in studying the differences among the categories or groups of the nominal variable in terms of a set of independent variables. It is my hope that researchers from any branch of social science will find the book useful for their research undertakings. Medical scientists may find it, particularly chapters on logistic regression and discriminant analysis, useful. Also, part 3 of the book may be quite useful to students of applied statistics. Although the book is meant mainly for no mathematically oriented readers, some mathematics needed to develop contexts for the techniques could not be avoided. Readers with no or little mathematical background can skip the mathematical sections and use the techniques without any problem since all techniques have been illustrated with examples; see, [15], [16], [17], [18], [19], [20] and [21].

#### REFERENCES

- [1] Anderson, T.W., (1958). An introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons, Inc.
- [2] Anscombe, F.J., (1973). Graphs in Statistical Analysis. *American Statistician*, 27, 17-21.
- [3] Campbell, D. T. & Stanley, J. C. (1963). *Experimental and Quasi-Experimental Designs for Research..* Chicago: Rand McNally.
- [4] Chatterjee, D.J. and Price, B., (1977). *Regression Analysis by Analysis*. New York: John Wiley and Sons.
- [5] Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: SAGE Publications.
- [6] Cronbach, L.J., (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- [7] Draper, N.R. and Smith, H. (1966). *Applied regression analysis*. New York, Wiley.
- [8] Durbin J. and Stuart, A., (1954). An experimental comparison between coders. *Journal of Marketing*, 19, 54-66.
- [9] Fisher, R. A. (1959). *Statistical Methods & Scientific Inference*. New York: Hafner Publishing.
- [10] Kish, L. (1965). *Survey sampling*. New York, Wiley.



- [11] Lachenbruch, P.A., (1975). Discriminant Analysis. New York, Hafner.
- [12] Leedy, P. & Ormrod, J. (2001). *Practical research: Planning and design* (7th ed.). Upper Saddle River, NJ: Merrill Prentice Hall. Thousand Oaks: SAGE Publications
- [13] Moser, C.A. and Kalton, G., (1971). Survey methods in social investigation. London, Heinemann Educational Books.
- [14] Pedhazur, E. J. (1997). Multiple Regression in Behavioral Research (Third Edition). Fort Worth: Harcourt Brace.
- [15] Saffir, M.A., (1937). A comparison study of scale constructed by three psychological methods. *Psychometrika*, 2, 179-198.
- [16] Solomon, R.L., (1949). An extension of control group design. *Psychological Bulletin*, 46, 137-150.
- [17] Student (1931). The Lanarkshire milk experiment. *Biometrika*, 23, 398-404.
- [18] Sufian, A.J.M (1984). The effects of infant and child Mortality on fertility in Bangladesh. Ph.D Dissertation. The Michigan State University, U. S.A.
- [19] Sufian, A. J. M. (1998). Methods and Techniques of Social Research. The University Press Limited, Dhaka, Bangladesh.
- [20] Tashakkori, A. & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: SAGE Publications.
- [21] 14. Tashakkori, A. & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social & behavioral research*. Thousand Oaks, CA: SAGE Publications.