CHAPTER 5
RESEARCH METHODOLOGY

5.1 INTRODUCTION

This chapter covers research design, determination of sample size, sampling design, questionnaire design, administration and structure of the questionnaire, scoring of the questionnaire, psychometric checks, reliability, validity, primary data, secondary data, period of the study, Framework of analysis and statistical tools used for the data analysis are presented here.

5.2 RESEARCH DESIGN

The study followed descriptive research design. It attempts to describe the impact of training programmes on employees’ organizational commitment such as affective commitment, continuous commitment and normative commitment. The stratified random sampling technique has been used to select the employees of TANGEDCO for the study.

5.3 DATA COLLECTION METHODS AND ITS TYPES

This study used both primary and secondary data. Survey method of data collection was adopted to collect the primary data. However the interview and observation method was adopted in an informal manner to reveal more interesting facts related to training and its impact on organizational commitment.
5.3.1 Primary Data

Data that have been collected through first-hand-experience is known as primary data. Primary data in this research are concerned with the survey instrument. A structured questionnaire was prepared for this purpose. The primary data was collected from the 1000 employees who are working in Tamil Nadu Generation and Distribution Corporation Limited (TANGEDCO). The designed questionnaire was circulated by the researcher through the trainers of the selected Training centres of Transmission and Distribution Training Institute, Madurai.

5.3.2 Secondary Data

The primary data was complemented by a spate of secondary sources of data. The secondary data pertaining to the study was gathered from the web portals of TNEB, Department of Government of Tamil Nadu, Tamil Nadu Industrial policy-2014, publications from TNEB Limited, Tamil Nadu Generation and Distribution Corporation Limited (TANGEDCO), Tamil Nadu Transmission Corporation Limited (TANTRANSCO) Energy Development Agency (TEDA), Electrical Inspectorate, Power Finance and Infrastructure Development Corporation Ltd, and Tamil Nadu Energy Development Agency (TEDA).

The above mentioned sources were very useful in writing introduction and industry profile chapters. Latest information was gathered from well-equipped libraries in Chennai, Coimbatore, Bangalore, and from internet web resources. The secondary data were used to identify the research gap through the literature survey from various national and international journals, magazines, periodicals, books and newspapers.
5.3.3 Period of the Study

The duration of the study was divided into four stages in a period of four years. In the first phase, the assortment of literature was done. In the second stage, the preparation of the questionnaire and its pre-testing with limited employees was accomplished. In the third period, the data collection, processing and analysis of data was done and in the final stage, the groundwork of the thesis was completed. The period of the study was confined to 2013-2016.

5.4 SAMPLING DESIGN

Sampling is a process of selecting a sufficient number of elements from a population. It increases the probability of results obtained from sample to be attributable to the population. For this research, the Tamil Nadu Generation and Distribution Corporation Limited (TANGEDCO) was designated for study.

5.4.1 Target Population

According to Sekaran & Bougie (2010), population is the entire group of people, events or things that researchers wish to investigate on. For the present study, Madurai district of Tamil Nadu State has been selected. So, in this research, target population is the total number of employees working in Tamil Nadu Electricity Board (TNEB).

5.4.2 Sampling Elements

The respondents for the study are those who are employed in Tamil Nadu Generation and Distribution Corporation Limited (TANGEDCO). Researcher has fairly targeted all the levels of employees from (TANGEDCO); in other words, the respondents consist of employees from all
hierarchy level (i.e. junior, middle, senior level), which include the employees working in technical and non-technical departments of Tamil Nadu Generation and Distribution Corporation Limited (TANGEDCO)

5.4.3 Sampling Technique

There are two types of sampling methods which are probability sampling and non-probability sampling. For probability sampling, each of the elements in the target population has an equal probability of being chosen as the sample for the survey conducted. Probability sampling is scientific, operationally convenient and simple in theory, and the results obtained from this method are more generalizable toward the target population. For non-probability sampling, each of the elements in the sampling frame does not have an equal chance to be chosen as the sample. Admittedly this method is simpler and convenient to operate however the results obtained cannot be confidently generalized to the population. One thousand sample respondents have been selected using stratified random sampling method from Tamil Nadu Generation and Distribution Corporation Limited where (n) 1000 samples have been collected for conducting the study.

5.4.4 Determination of Sampling Size

A population is defined as the “total collection of individuals or objects that forms the focus of the research” whereas the sample is “a selected part or a subset of the population (Pretorius 1995). According to Pretorius (1995), research is generally conducted to make inferences about the population based on the information available about the sample, in order to make inferences from the sample to the population.
A number of formulae have been formulated for determining the sample size depending upon the availability of information. The researcher has used the below mentioned formulae for calculation of sample size for an unknown population.

\[
\text{Sample size } n = \left( \frac{ZS}{E} \right)^2
\]

where,  
\( Z \) = Standardized value corresponding to a confidence level of 95% = 1.96  
\( S \) = Sample SD from pilot study of 50 sample = 0.8066  
\( E \) = Acceptable error 5% = 0.05

Hence

\[
\text{Sample size } n = \left( \frac{1.96 \times 0.8066}{0.05} \right)^2 = 999.7435 \approx 1000
\]

Where \( n = 1000 \)

5.5 RESEARCH INSTRUMENT

5.5.1 Questionnaire Design

The researcher struggled by all means to identify the appropriate structured questionnaire developed by the eminent researchers in the preferred research topic. Even though there are vast structured questionnaires accessible in the preferred research area, none of the structured questionnaires suited the chosen research topic context in Tamil Nadu Electricity Board (TNEB). i.e. the impact of training and development programmes on the employees of...
Tamil Nadu Electricity Board. Most of the structured questionnaires accessible had only very few questions related to Tamil Nadu Electricity Board. Henceforth the investigator has designed a new investigation instrument and also confirmed the reliability, validity and content validity of the designed questionnaire after the pilot study.

5.5.2 Administration of the Questionnaire

The questionnaire can be administered individually or in group. In order to make the Tamil Nadu Generation and Distribution Corporation Limited employees feel free, the supervisor/manager’s presence has been kept away. Moreover, the respondents should remain incognito. This has given them a greater sense of security. The time generally taken for completion of one questionnaire has been thirty minutes. The purpose of the questionnaire is to measure the impact of training and development programmes and organizational commitment of employees towards TANGEDCO after attending the training programmes in Tamil Nadu Generation and Distribution Corporation Limited (TANGEDCO).

5.5.3 Pilot Study

The questionnaire prepared for the respondents was pre-tested with 50 employees who are working in Tamil Nadu Generation and Distribution Corporation Limited (TANGEDCO) and taken training recently at Transmission and Distribution Training Institute, Madurai. After pre-testing, required amendments were made in the questionnaire to fit in the path of the current study. Finally, the questionnaire was checked for the reliability test with fifty samples. As per the assessment test for reliability, Cronbach’s alpha values divulge a high scale reliability alpha = 0.962 for the segment relating to impact of training and development programmes, alpha = 0.981 for the
segment of affective commitment and 0.929 for the segment of continuance commitment and alpha = 0.973 for normative commitment.

5.6 CONSTRUCTS MEASUREMENT (SCALE AND OPERATIONAL DEFINITIONS)

5.6.1 Scoring of the Questionnaire

The scale against which the respondents have indicated the extent of agreement / disagreement with reference to the characteristics of his/her organization is defined by the following five categories.

- Strongly agree - 5
- Agree - 4
- Neutral - 3
- Disagree - 2
- Strongly disagree - 1

To find out the raw scores for each employee, the scores of all items in the questionnaire answered by him/her have been added. This has been given as the score of that particular employee regarding his/her evaluated impact of training and development programmes attended and employees’ organizational commitment after attending the training and development programmes.

5.6.2 Scaling Technique

Nominal scale and Likert’s scale have been used in this study when the questionnaire has been developed.
5.6.2.1 Nominal scale

Nominal scale is simply a system of assigning number symbols to events in order to label them. Nominal scales provide convenient ways of keeping track of people, objects and events. This scale is used for the demographic and training details section of the questionnaire where the questions are categorized variables.

5.6.2.2 Likert`s scale

In a Likert`s scale, the respondent is asked to respond to each of the statements in terms of several degrees, usually five degrees of agreement or disagreement. At one extreme of the scale there is strong agreement with the given statement and at the other, strong disagreement, and between them lie immediate points. Five point Likert`s scale (5- Strongly Agree, 4- Agree, 3- Neutral, 2-Disagree, 1- Strongly Disagree) has been used for all dimensions except demographic profile and training details section which consists of their own choice.

5.7 DATA PROCESSING

After collecting the questionnaires from the respondents, data processing step is to be taken before any analysis is being implemented. The data preparation process consists of checking, editing, coding, and transcribing. Meanwhile, all the unusual responses are identified.

Before checking all the collected questionnaires, every questionnaire has been counted and numbered to assure that the questionnaires distributed have been returned by the respondents. Then it is ensured that all the questions in the questionnaire have been filled up completely by the respondents without any omission. Incomplete
questionnaires have been removed. Coding, data entry, editing and data transformation for all the remaining survey questionnaires have been done.

5.7.1 Coding

This step is taken in order to key the data into the Statistical Package for Social Sciences (SPSS) system. Code has been assigned to each participant’s response. For instance, in Section A-, under the Personal Particulars heading where the department of work has to be specified the researcher has assigned “1” to technical department and “2” for non-technical. While for the question of work experience, the researcher has assigned “1” to “3” for all the responses.

Apart from that, the researcher attributed “1” for male and “2” for female. For the questions in Section B, D and E, the code has been marked from “1” to “5” for all the responses, where “1” symbolizes strongly disagree and “5” symbolizes strongly agree.

5.7.2 Data Entry (Transcribing) and Editing

Once all the questions’ responses have eventually been coded, the researcher has entered all the data into the SPSS database. This process is also known as transcribing (Malhotra 1993). Before running the reliability test, the researcher has carried out the editing tasks towards all the responses. The researcher has attempted to detect and correct the problems, such as illogical, inconsistent or illegal responses.

Illogical response is the response that is given by the respondent which looks significantly different from others’ responses. Sometimes, this respondent is known as outlier. While, inconsistent responses happen when the respondents’ responses is incoherent with other information provided.
Also, it is possible that the inconsistent responses are caused by bias. As a result, the researcher needs to edit the inconsistent responses provided by the respondents.

### 5.7.3 Data Transformation

Additionally, the researcher has also carried out data transformation after data entry and editing. Data transformation is a data coding variation, which is the process of altering the original numerical representation of a quantitative value to another value (Sekaran and Bougie, 2010). The data transformation is not required in this research, because all the questions are in positive forms.

### 5.8 DATA ANALYSIS

According to Zikmund et al. (2010), data analysis is defined as the reasoning application which helps the researchers to understand the data that have been collected. The purpose to implement data analysis is to examine and model the data by assigning facts and figure to answer research problem. Also, it highlights the useful information by recommending assumptions to take advantage of the collected data in order to solve some specific problem, such as addressing the research problem.

The computer software that has been applied to analyze the collected data is IBM SPSS 22.0. IBM SPSS 22.0 provides us many types of analysis that is very helpful in our current research. Typically, there are three types of analysis that is required in our research, which are descriptive analysis, scale measurement and inferential analysis.
The core of the study is the impact of training and development programmes on the employees’ organizational commitment in Tamil Nadu Generation and Distribution Corporation Limited (TANGEDCO), hence the study centers on the independent variable ‘impact of training and development programmes’ and dependent variable viz., employee organizational commitment and its relationship with the related independent variables such as affective commitment, continuous commitment and normative commitment.

### 5.8.1 Descriptive Analysis

Descriptive analysis is used in order to clarify and describe the characteristics of the variables of interest in a situation (Sekaran and Bougie, 2010). Besides, Zikmund defined descriptive analysis as the elementary transformation of data in a way that illustrate the fundamental characteristic, such as central tendency and variability. Generally, mean, median, mode, variance, range and standard deviation are widely applied in describing the descriptive statistics. The advantage of using descriptive analysis is that it helps to summaries the sample and measure. It also forms basic quantitative data analysis with simple graphics analysis.

In this research, Descriptive analyses have been done for the responses derived from ‘Section A’ Personal details of the respondents by using frequency analysis with frequency and percentage.

### 5.8.2 Scale Measurement- Reliability Test

#### 5.8.2.1 Psychometric checks

As mentioned earlier, a structured questionnaire developed by the researcher has been used as the instrument for data collection for the study. Items selected for the constructs have been mainly adopted from prior studies.
to ensure content validity. However, the instrument has been validated for the main study with a size of 1000 respondents.

5.8.2.2 Reliability

Reliability, also called consistency and reproducibility, is defined in general as the extent to which a measure, procedure, or instrument yields the same result on repeated trials (Carmines & Zeller, 1979). It can be used to assess the degree of consistence among multiple measurements of variables (Hair, Anderson, Tathman, & Black, 1998). The internal reliability of the measurement models has been tested using Cronbach’s alpha and Fornell’s composite reliability (Fornell and Larcker 1981). The Cronbach’s reliability coefficients of all variables should be higher than the minimum cut-off score of 0.70 (Nunnally 1978; Nunnally and Bernstein, 1994). The questionnaire meant for the respondents has been pre-tested with 50 employees from Tamil Nadu Generation and Distribution Corporation Limited (TANGEDCO) at Madurai. After pre-testing, necessary modifications have been made in the questionnaire to fit in the track of the present study. Finally, the questionnaire has been checked by the reliability test for the fifty samples. As per the assessment test for reliability, Cronbach’s alpha values divulge a high scale reliability alpha = 0.962 for the segment relating to impact of training and development programmes, alpha = 0.981 for the segment of affective commitment and 0.929 for the segment of continuous commitment and alpha = 0.973 for normative commitment.

5.8.2.3 Validity

A scale is said to be valid if it measures correctly what it is expected to measure. In other words, a scale is valid only when it is real and correct. The validity of a questionnaire relies first and foremost on reliability. If the questionnaire cannot be shown to be reliable, there is no discussion of
its validity. Researchers use different methods of establishing the validity of the instrument which they have developed. They are: content validity, convergent validity, discriminate validity and nomological validity. In the present study the content validity is established. It is given in the following section.

5.8.2.4  Content Validity

For the content validity, a thorough review of the literature has been conducted. As mentioned earlier, all items of the constructs have been drawn from well-established studies to ensure content validity. The questionnaire has also been reviewed by a panel of expert’s i.e. senior manufacturing professionals and human resource managers working in the manufacturing sector and academicians. The changes suggested by the panel members have been incorporated to improve both the content and clarity of the questionnaire. The instrument has been tested through two stages. In the first stage, two English faculty members have reviewed the instrument to ensure the clarity of items and the accuracy of the language. In the second stage, a panel of experts has been selected to establish face and content validity of the instrument. The panel of experts consisted of six individuals, two senior manufacturing professionals, two human resource managers of the manufacturing companies, who had earlier participated in the instrument development and two senior academicians.

5.9  STATISTICAL TOOLS FOR DATA ANALYSIS

The requirement and importance of statistics is escalating, especially in social sciences and management research. It is important to recognize an appropriate statistical design which brings solutions to the entire research hypotheses. Statistics are the tools used to check our facts about the data.
Tools for data analysis includes

- Frequency distribution
- Chi square
- Independent Sample ‘t’ test
- One way ANOVA
- Kalmogorov-Smirnov test
- Friedman test
- Correlation
- Multiple regression analysis
- Factor analysis
- Structural equation modeling

5.10.1 Frequency Distribution

A frequency table is a simple way to display the number of occurrences of a particular value or characteristic.

5.10.2 Chi-Square test

A Chi-square is a statistical measure used in the context of sampling analysis for comparing a variance to a theoretical variance. As a non-parametric test, it can be used to determine if categorical data shows dependency or the two classifications are independent. It can also be used to make comparisons between theoretical populations and actual data when categories are used. Thus, the chi-square test is applicable in large number of problems. The test is, in fact, a technique through the use of which it is possible for all researchers to (1) test the goodness of fit (2) test the
significance of association between two attributes, and (3) test the homogeneity or the significance of population variance.

**5.10.3 Independent Samples T Test**

The independent samples t test allows the researcher to evaluate the mean difference between two populations using the data from two samples. This test is used in situations where a researcher has no prior knowledge about either of the two populations being compared. The general purpose of the independent samples t test is to determine whether the sample mean difference obtained is a real difference between the two populations or simply the result of sampling error.

**5.10.4 One-Way ANOVA**

Analysis of variance procedures are powerful parametric methods for testing the significance of differences between sample means where more than two conditions are used, or even when several independent variables are involved. ANOVA makes it feasible to appraise the separate or combined influences of several independent variables on the experimental criterion (Mouton & Marais 1990). ANOVA test has been therefore used to identify whether there is a statistical significant difference between the demographical variables and impact of health and safety training programmes, employees work related attitude.

**5.10.5 Kolmogorov-Smirnov Test**

The Kolmogorov-Smirnov test (Chakravart et al. 1967) is used to decide if a sample comes from a population with a specific distribution. The Kolmogorov-Smirnov (K-S) test is based on the empirical distribution
function (ECDF). Given N ordered data points Y1, Y2, ..., YN, the ECDF is defined as

\[ EN = \frac{n(i)}{N} \]

where \( n(i) \) is the number of points less than \( Y_i \) and the \( Y_i \) are ordered from smallest to largest value. This is a step function that increases by \( 1/N \) at the value of each ordered data point.

An attractive feature of this test is that the distribution of the K-S test statistic itself does not depend on the underlying cumulative distribution function being tested. Another advantage is that it is an exact test (the chi-square goodness-of-fit test depends on an adequate sample size for the approximations to be valid).

5.10.5 The Friedman Test

The Friedman test is a test for comparing three or more related samples and which makes no assumptions about the underlying distribution of the data. The data are set out in a table comprising \( n \) rows by \( k \) columns. The data are then ranked across the rows and the mean rank for each column is compared.

5.10.6 Correlation

Degree and type of relationship between any two or more quantities (variables) in which they vary together over a period; for example, variation in the level of expenditure or savings with variation in the level of income. A positive correlation exists where the high values of one variable are associated with the high values of the other variable(s). A 'negative correlation' means association of high values of one with the low values of the other(s). Correlation can vary from +1 to -1. Values close to +1 indicate a high-degree
of positive correlation, and values close to -1 indicate a high degree of negative correlation. Values close to zero indicate poor correlation of either kind, and 0 indicates no correlation at all. While correlation is useful in discovering possible connections between variables, it does not prove or disprove any cause-and-effect (causal) relationships between them.

5.10.7 Multiple Regression Analysis

Multiple regression analysis the common and separate influences of two or more variables on a dependent variable (Kerlinger 1986), and it is used to establish the extent to which various differing variables add to predict another variable (Guyatt et al. 1995). Multiple regression has been therefore used to study the dependent variable (Employee’s Performance) is statistically significant on the variance in independent variables such as organizational commitment, overall impact of health and safety training and job satisfaction.

5.10.8 Factor Analysis

A factor is an underlying dimension that accounts for several observed variables. There can be one or more factors, depending upon the nature of the study and the number of variables involved in it. Factor analysis involves many terminologies which are presented in this subsection for better understanding of the related techniques.

Correlation coefficients matrix is the original observations between different pairs of input variables. Factor loadings matrix representing the correlation between different combinations of variables and factors. Communality is the sum of squares of the factor loadings of the variable ‘i’ on all factors. Eigen value is the sum of squares of the factor loadings of all variables on a factor.
After obtaining factor loadings, one should examine whether the factor loading matrix possesses a simple structure. If a factor loading matrix has a simple structure, it is easy to make interpretations about the factors. If there is no simple structure, then the n – dimensional space of the factors should be rotated by an angle such that the factor loadings are revised to have a simple structure which will simplify the process of interpretation of the factors. Such rotation is called rotation of factors. A simple structure means that each variable has a very high factor loading (as high as 1) on one of the factors and very low factor loading (as low as 0) on other factors. The communalities of each variable before and after factor rotation will be the same. The popular methods of rotation of factors are varimax method and promax method. Varimax method of factor rotation employs orthogonality between different pairs of factors axes. This means that the angles between different pairs of factors axes are 90° even after rotation. The promax method employs oblique rotation. This means that the angles between different pairs of factors axes are not 90° after rotation. Both the techniques aim at better interpretations.

5.10.9 Structural Equation Modeling (SEM)

An explanation for the use of Structural equation modeling with AMOS and methods to assess construct validity and reliability for all measures is addressed in this study. The research or model describes the causal relationship among reaction to training programmes, skills acquisition, behavioural change, effect of training, organizational commitment, employee’s performance and job satisfaction. These paths are related to causal processes. Thus the Structural Equation Modeling (SEM) approach is necessary in order to examine these variables. The data analysis has been carried out by means of SPSS (statistical package for the social science,
Structural Equation Model (SEM) with AMOS 20 software provided several indicators to assess fit. The confirmatory factor analysis has showed the acceptable model fit by including the normal fit index (NFI), the comparative fit index (CFI), the root mean square error of approximation (RMSEA), hypothesis model (Research model) based on the research hypothesis and review of the theoretical and empirical literature, the hypothesis model has been examined in this study. There are two latent variables; impact of health and safety training programmes and employees’ work related attitude to explore the cause and effect relationship among these variables in the hypothesized model.

SEM is a statistical methodology with a confirmatory approach to analyze multivariate data (Byrne 2001). The general SEM model is composed of two sub models; a measurement model and a structural model. James et al. (1982) recommended the measurement model testing first, followed by full structural model testing.

Statistical significance for all analysis has been set at less than 0.05. The measurement model identifies relationships between the observed and latent variables. By means of CFA, the measurement model provides the link between scores on an instrument and the constructs that they are designed to measure. Hence structural model identifies the causal relationships among the latent variables and specify that particular latent variables directly or indirectly influence certain other latent variables in the model (Byrne 2001).