CHAPTER - III

DATA BASE AND RESEARCH METHODOLOGY

The basic purpose of research is to find out solutions to the predetermined questions by making use of the scientific and systematic techniques. Before finding a solution to the problem, we design an approach in which we want to proceed in future. This refers to the development of an appropriate research design. Research design is concerned with the methods and ways in which the investigator manages the situation to study the selected problem. “A research design is the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure” (Jaboda et al., 1952). In other words, research design is a process of deliberate application of research methods directed towards bringing an expected situation under control.

3.1 POPULATION/UNIVERSE AND SAMPLE

According to Creswell (2005), the term population is defined as ‘a group of persons having a common characteristic, for instance, all secondary school teachers or all secondary school principals in a particular district, region or country would form a population’. The whole subject matter from which a portion is taken is technically termed as ‘Universe’ or ‘Population’ and the portion selected for the research purpose is known as ‘sample’. ‘A sample is a small representation of the large whole’ (Young, 1951). In any research, it is difficult to deal with the whole universe. In order to save time, money and efforts sampling is done. Purposive and convenience random sampling method has been used to select the sample for the present study. The sampling was done at two levels. Firstly, the schools were selected purposely, while, the sample of teachers were drawn on the basis of convenience random sampling. The choice of selecting the private unaided public schools was done because of the following reasons:

1. No comprehensive research has been conducted on private un-aided public schools in Punjab; and
2. Number of public schools in Punjab has grown in number during the last two decades.

The universe of the study is comprised of the unaided public school teachers in the state of Punjab. The study is based on primary and secondary sources of data. The primary data was collected from the schools located in three regions i.e. Majha, Doaba and Malwa of Punjab. One district from each region namely, Amritsar, Jalandhar and Ludhiana respectively, was selected for the sample. The following categories of co-educational schools (both urban and rural) were covered for the collection of data:

1. Schools managed by D.A.V. Management Committee;
2. Schools managed by Chief Khalsa Dewan (C.K.D.) Committee and other Sikh institutions; and

Out of these management committees following schools were selected:

| Table 3.1 |
|---|---|
| Region and Organisation- Wise list of Schools |
| Region | Organisation |
| Majha (Amritsar Dist.) |
| 1 | Sri –Guru Harkrishan Senior Secondary School, Amritsar | C.K.D. |
| 2 | Sri –Guru Harkrishan Public School, Ajnala | C.K.D. |
| 3 | DAV Public School, Lawarence Road, Amritsar | D.A.V. |
| 4 | DAV Public School, Bhikhiwind | D.A.V. |
| 5 | St. Francis School, Amritsar | C.M. |
| 6 | St. Francis School, Jandiala Guru | C.M. |
Doaba (Jalandhar Dist.)

<table>
<thead>
<tr>
<th>No.</th>
<th>School Name</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>MGN Public School, Jalandhar City</td>
<td>Others</td>
</tr>
<tr>
<td>8</td>
<td>Guru Nanak Public School, Nakodar</td>
<td>Others</td>
</tr>
<tr>
<td>9</td>
<td>Daya Nand Model DAV Public School, Jalandhar</td>
<td>D.A.V.</td>
</tr>
<tr>
<td>10</td>
<td>D R DAV Public School, Phillour</td>
<td>D.A.V.</td>
</tr>
<tr>
<td>11</td>
<td>St. Joseph’s Convent School, Goraya</td>
<td>C.M.</td>
</tr>
<tr>
<td>12</td>
<td>St. Joseph’s Public School, Jalandhar</td>
<td>C.M.</td>
</tr>
</tbody>
</table>

Malwa (Ludhiana Dist.)

<table>
<thead>
<tr>
<th>No.</th>
<th>School Name</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>Guru Nanak Public School, Ludhiana</td>
<td>Others</td>
</tr>
<tr>
<td>14</td>
<td>Sri-Guru Harkrishan Public School, Sidhwan Khurd</td>
<td>C.K.D.</td>
</tr>
<tr>
<td>15</td>
<td>DAV Public School, Jagraon</td>
<td>D.A.V.</td>
</tr>
<tr>
<td>16</td>
<td>DAV Public School, Ludhiana</td>
<td>D.A.V.</td>
</tr>
<tr>
<td>17</td>
<td>Sacred Heart Senior Secondary School, Ludhiana</td>
<td>C.M.</td>
</tr>
<tr>
<td>18</td>
<td>Sacred Heart Convent Senior Secondary School, Khanna</td>
<td>C.M.</td>
</tr>
</tbody>
</table>

Source: Primary Data

Further, sample, of 300 teachers belonging to three different categories i.e. Primary Teachers (PRT), Trained Graduate Teachers (TGT), and Post-Graduate Teachers (PGT), was planned. Convenience random sampling method was employed for the selection of the respondents from the selected schools. At the initial stage 300 teachers were planned. However, only 272 respondents’ responses were taken for final analysis after the scrutiny of research schedules. The region wise distribution of the total respondents was Majha(90), Doaba(86) and Malwa(96). On the other hand the
organization wise number of interviewed teachers were: Sikh managed institutions (91), Hindu managed institutions (98) and Christian managed institutions (83). The data for the study were collected with the help of a comprehensive research schedule (structured), which included all important aspects of their socio-economic profile and job related factors. Majority of the questions in the research schedule were close ended, whereas, a few questions were open ended. A set of statements, was taken from the existing literature relating to the factors which influence the level of job satisfaction. All the respondents were interviewed by the researcher herself. A five point Likert Type Scale was applied to record the perception of respondents on the various issues. The secondary data was collected from secondary sources like books, journals and other published/unpublished reports, internet etc.

3.2 DATA ANALYSIS

All the interview schedules were scrutinized and incomplete schedules were excluded from the analysis. Further, the research schedules were coded to enter the data in the computer. Commensurate with the different objectives of the study, various tools of analysis have been used. Statistical Package for Social Sciences (SPSS) version 17.0 for windows was used to perform the statistical analysis. The analysis of data was carried out with the help of various statistical tools like Frequency Distribution, Weighted Average Score, Cross Tabulation and Factor Analysis. Cross Tabulation and Factor Analysis, are the basic techniques that provide rich insights into the data and lay the foundation for more advanced analysis (Naresh K. Malhotra and Satyabhushan Dash, 2009).

3.3 DESCRIPTION OF ANALYTICAL TOOLS

TABULAR ANALYSIS

Tabulation is the process of considering data in the form of a table, so that it may be more easily understood, and comparison if any be made. The main purpose of tabular analysis is to summarise the numerical information in a very simple form along with a brief explanation. Tabular analysis was carried out for the following purpose:
1. To examine the socio-economic profile of the teachers;
2. To study the perception of teachers; and
3. To examine the Region and Organization (Religion wise) profile of the teachers.

**WEIGHTED AVERAGE SCORE**

Majority of responses were recorded on five point Likert Type Scale and Weighted Average Scores (WAS) were calculated. The weights were assigned to five levels; namely, strongly agree to strongly disagree in the order of 5, 4, 3, 2, and 1. Due care was given while assigning weights to the different statements. Negative statements were adjusted accordingly. Consequently, weighted average scores were calculated on the different parameters to know the level of satisfaction and awareness among the teachers working in the sampled schools.

**CROSS TABULATION**

To determine the level of satisfaction of the teachers on the basis of selected personal, demographic and job characteristics, Cross Tabulation analysis has been used. Cross tabulation is applicable to all types of categorical variables: one in which the different categories of variables cannot be quantified, such as marital status, gender, qualifications etc. and one in which the different categories are associated with quantifiable number, such as age, income, experience etc. or imprecisely quantifiable numbers such as respondents agreement or disagreement with a particular attitude, (Boyd et al. 1990). Cross Tabulation is a technique that describes two or more variables simultaneously and results in tables that reflects the joint distribution of two or more variables that have a limited number of categories or distributions. A frequency distribution describes one variable at a time, whereas a Cross Tabulation describes two or more variables simultaneously. It helps us to understand how one variable such as brand loyalty relates to another variable such as sex.

A cross tabulation includes a cell for every combination of the categories of the two variables. The number in each cell shows how many respondents gave that combination of responses. Cross Tabulation with two variables is also known as
Bivariate Cross Tabulation. It is important to note that Cross Tabulation examines association between variables not causation. The chi-square statistics ($\chi^2$) is used to test the statistical significance of the observed observations in cross-tabulation. It assists us in determining whether a systematic association exists between the two variables or not. To determine whether a systematic association exists, the probability of obtaining a value of chi-square as large as or larger than the one calculated from the Cross Tabulation is estimated.

For this purpose, the two levels of satisfaction, i.e. high and low were taken as dependent variables while other variables were taken as independent variables for the present analysis. It may be mentioned here that two levels of satisfaction were generated as follows: we first find the weighted average score in respect to five statements relating to different attributes of job satisfaction for each response measured on a five point Likert Type Scale. These five statements are selected to calculate the high and low level of job satisfaction including the following: satisfaction in regard to earning; housing facilities; job security; promotion opportunities and Izzat/status. These weighted response score were then aggregated across all the respondents in the sample and the average score was determined. The respondents experiencing score higher than the average (so computed) are classified as having high level of satisfaction ,whereas the ones having score less than or equal to the average are classified as having low level of satisfaction. The value of contingency co-efficient is used to test the level of significance of the variables. In other words, an attempt was made to find whether the relationship among the related variables is significant or not.

**RELIABILITY OF SCALES**

Cronbach’s Alpha test was used to check to reliability of the scale used for the collection of data. The value of Cronbach’s Alpha was 0.786. The alpha value is inflated by large number of variables, so there is no set interpretation, as to what is an acceptable alpha value. A rule of thumb that applies to most situations is that if the value of $\alpha$ is $>0.7$, it is considered as acceptable, which means that the scale is reliable (George and Mallory, 2011).
FACTOR ANALYSIS

The general purpose of factor analysis is to find a way in condensing (summarizing) the information contained in number of original variables into a smaller set of net, composite dimensions (factors) with a minimum loss of information, i.e., to search for and define the fundamental constructs or dimensions assumed to underline the original variables (Joseph et al., 1995).

To identify the factors which influence the job satisfaction of the teachers, Factor Analytical Approach has been used. This is a statistical approach that can be used to analyse interrelations among a large number of variables and to explain these variables in terms of their common underlying dimensions (factors). The factor analysis is designated as the queen of analytical methods because of its power and elegance (Diwedi, 1997).

The Factor Analysis is the most frequently used tool to identify a small number of factors (e.g., outgoing) that may be used to represent a relationships among sets of interrelated variables. Cattell’s use of Factor Analysis underlines its primary usefulness, i.e., to take a large number of observable instances to measure unobservable construct or constructs (George, Darren and Mallery, Paul, 2011).

KEY STATISTICS ASSOCIATED WITH FACTOR ANALYSIS

Before proceeding for the further analysis in factor analysis the following terms have been explained to understand the correlation matrix generated on the basis of the primary data with the help of Statistical Package for Social Sciences (SPSS) software which is a powerful tool capable of conducting any type of data analysis used in the social sciences, natural sciences or in the business world. The following terms are the frequently used for the interpretation of the factor analysis solution. These are Bartlett’s Test of Sphericity, Kaiser-Meyer-Olkin, Communality, Eigenvalue, Percentage of Variance, Correlation matrix etc.

Bartlett’s test of Sphericity

Bartlett’s test of Sphericity is a statistics used to examine the hypothesis that the variables are uncorrelated in the population. In other words, the population correlation
matrix is an identity matrix; each variable correlates perfectly with itself (r=1) but has no correlation with the other variables (r=0).

**Kaiser-Meyer-Olkin**

KMO measure of adequacy is an index used to examine the appropriateness of factor analysis. A high value of KMO (between 0.5 and 1) indicates that factor analysis is appropriate. Value below 0.5 implies that factor analysis may not be appropriate.

**Correlation matrix**

A correlation matrix is a lower triangle matrix showing the simple correlations, between all possible pairs of variables included in the analysis. The diagonal elements, which are all 1, are usually omitted.

**Factor loading plot**

It is the plot of the original variables using the factor loadings as the coordinates.

**Factor matrix**

It contains the factor loading of all the variables on all the factors.

**Factor scores**

They are the composite scores estimated for each respondent on the derived factors.

**Communality**

It is the amount of variance a variable shares with all the other variables being considered. This is also the proportion of variance explained by the common factors.

**Eigenvalue**

The Eigenvalue represents the total variance explained by each factor.

**Percentage of variance**

This is the percentage of the total variance attributed to each factor.

Factors are underlying constructs that describe a set of variables. With factor analysis the procedure is similar, to a forward run in multiple regression. However, factor analysis does not begin with a dependent variable. The first step in factor analysis is for the computer to select the combination of variables whose shared correlation
explains the greatest amount of total variance. This is called factor I (or component I, the words are used interchangeably). The factor analysis will then extracts a second factor. This is the combination of variables that explains the greatest amount of the variance that remains, i.e., variation after the first factor has been extracted. This is called factor II or component II. This procedure continues for a third factor, fourth factor and so on, until as many factors have been extracted as there are variables. The steps involved in conducting the Factor analysis are presented in the Figure 3.3.1.

**Figure 3.3.1**

**Factor Analysis**

- Formulate the Problem
- Construct the correlation matrix
- Determine the method of factor analysis
- Determine the number of factors
- Rotate the factors
- Interpret the factors
- Calculate the factor
- Select the surrogate factor
- Determine the model fit

Source: Hair et al. (1995)
The first step is formulating the problem, where the variables to be included should be specified on the basis of past research, theory and judgment of the researcher. The second step is the calculation of correlation matrix of all variables of interest. If the correlations between all the variables are small, factor analysis may not be appropriate. The third step is to determine the method of Factor analysis. There are two basic models that the analyst can utilize to obtain factor solutions. They are known as Common Factor and Principal Components Analysis. Both the models are widely used. Selection of the extraction method depends upon the analysts’ objective. Principal Component Analysis is used when the objective is to summarise most of the original information (variance) in a minimum number of factors for prediction purposes. In contrast, Common Factor Analysis is used primarily to identify underlying factors or dimensions reflecting what the variables share in common (Hair et al., 1995). The fourth step is to determine the number of factors. In order to summarize the information contained in the original variables, a smaller number of factors should be extracted. Many procedures are suggested for determining the number of factors to be extracted. The following criterion is used to determine the number of factors:

1. In a Priori Criterion, the analyst already knows how many factors to extract and accordingly instructs the computer.

2. In Latent Root Criterion, only those factors which have latent roots greater than 1 are considered significant.

3. In Percentage of Variance Criterion, the cumulative percentage of variance extracted by successive factors is considered. In social sciences it is common to consider a solution satisfactory when it accounts for more than 60 percent of the total variance (and sometimes even less).

4. In Scree Test Criterion, at least one factor more than latent root criterion is usually extracted. The later factors extracted in principal component factor analysis model, contain both common and unique variance- the proportion of unique variance is much higher in later than in earlier factors.
5. Determination based on Eigenvalues, the factors having Eigenvalues more than one are retained.

In Percentage of variance, the number of factors extracted is determined so that the cumulative percentage of variance reaches a satisfactory level. In the present case the number of factors extracted was decided on the basis of criterion 3 and 5 i.e. cumulative percentage of variance and the value of eigenvalue.

The next step in factor analysis is to rotate the factors which results in factor matrix. The factor matrix contains the coefficients used to express the standardized variables in terms of the factors. These coefficients, the factor loadings, represent the correlation between the factors and the variables.

Loading are rotated to make them more interpretable by making the loadings for each factor either large or small, not in between. For rotation either Orthogonal or Oblique method can be employed. In the present case Orthogonal rotation along with the Varimax method of rotation of factors was used in order to have more clarity in factor solution. It is one of the most popular Orthogonal Rotation Procedure. ‘The Varimax criteria maximize the sum of the variances of the squared loadings within each column of the loading matrix. This tends to produce some high loadings and some loadings near zero, which is one of the aspects of simple structure’ (Dunteman, 1989).

The next step is the interpretation of the factors, by identifying the variables that have large loadings on the same factor. The factor can be interpreted in terms of the variables that load high on it. The correlation between the variables can be attributed to common factors. The differences between the observed correlations (as given in the input of correlation matrix) and the reproduced correlations (as estimated from the factor matrix) can be examined to determine the model fit.

3.4 NULL HYPOTHESIS

The following null hypothesis has also been tested on the basis of findings of the study.
1. Ho: There is no significant difference in the level of job satisfaction between male and female teachers in the un-aided public schools;

2. Ho: Income has no influence on the level of job satisfaction of teachers;

3. Ho: Type of the family does not influence the level of job satisfaction of teachers;

4. Ho: Length of service has no influence on the level of job satisfaction;

5. Ho: Age has no impact on the level of job satisfaction of school teachers;

6. Ho: There is no significant difference in satisfaction level of graduate teachers and post graduate teachers in these schools;

7. Ho: The designation of teacher has no influence on the level of job satisfaction.

3.5 LIMITATIONS OF THE STUDY

One of the respondents said that the level of job satisfaction changes every day. The present study was limited by the following-

1. The present study was confined to an analysis of un-aided public school teachers in Punjab. Therefore the results may not generalize to teachers employed in other settings of the other states.

2. The tool used to collect the data was the research schedule. There might be some differences in the perception of the respondents and what is actually recorded.

3. The ability to generalize from the findings collected in a short time period in which the study took place may also be a limitation.

4. Management’s response regarding the policies of the school could not be studied which could have provided an important insight to the level of job satisfaction of the teachers.

5. Data collection was done by quantitative means only. This procedure does not seem to capture the complexity of teachers’ perceptions of their workplace conditions. Therefore, a combination of quantitative and qualitative research
would have been a better option. Interviews could have assisted in achieving a better assessment of the participants’ lives at work and a better indication of the exact factors that contributed to their levels of job satisfaction and dissatisfaction.

6. In this study the questionnaires were completed during the examination period, which is usually a stressful time for teachers. This may have had an impact on the responses of the teachers. Since they were preoccupied with duties related to the examinations, their satisfaction level could have been different, as compared to their normal routine working.

7. Efforts were made to collect data from both rural and urban schools, however, majority of the schools were located in urban areas only. As result of which rural–urban comparison in regard to job satisfaction was not possible. Similarly, due to higher share of female teachers in all schools, more respondents in the sample were female teachers; therefore, the result of the study might have influence of gender biasness.