CHAPTER 4

RESEARCH METHODOLOGY

4.1 INTRODUCTION

Research methodology plays an important role in any research. It includes research design, data collection, analysis and interpretation of results. The term ‘methodology’ comprises this whole process. The final results of a research depend on the methodology that we are employing and methodology depends on the type of data needed to answer the research questions. Social science researchers use either quantitative research methods, qualitative research methods or both (triangulation). But there is a controversial argument between social science researchers and scientific researchers regarding the use of these two methods in the social sciences and debates on quantitative and qualitative research methods is still continuing. However, these two approaches help to understand the socio-economic realities of the society. The qualitative approach uses non-numeric data and the quantitative approach uses numeric data. Selecting an appropriate methodology for specific research depends on the research objectives and research questions which are to be answered through the research. The research methods must lead to comprehensive and clear results at the end of the research.

The questionnaire is a measuring tool (Oppenheim 1992) loosely; a questionnaire consists of a series of questions, checklists, attitude scales and a variety of other approaches in a structured sequence. They are used to provide
descriptive and or analytical information which is suitable for statistical analysis.

Questionnaires usually involve large samples and are costly so it is essential to plan the research approach. Who to question, types of questions to ask, sample size, inherent biases, and these are amongst the factors that affect questionnaire measurement, specification and procedures (Oppenheim 1992). Questionnaires needs exploratory work, design and planning before any specification can be established. A certain rigidity of questioning and sampling procedure is needed to maintain the statistical validity which makes them relatively inflexible. Questionnaire assessments of consumption of fuel wood, fodder, and food grains are, for example, indirect unless a weighting measurement is included and are dependent on the accuracy of recall by the respondent (Oppenheim 1992).

The methodology used in the study is an integrated methodology, where traditional schedule based data collection and processing is integrated with the modern, statistical as well as qualitative analysis. The former complements the latter. The methodology which follows the traditions of social science research (Kundu 1992, Wood 1996) and the latest developments in economic research have the following components:

1. Field survey (primary data).
2. Collection of documented data (secondary data).
3. Statistical approach
4. Analysis and interpretation of Teachers’ and Parents’ data.

In selecting the most appropriate tool, the following considerations were useful: the Uses, Resources, Familiarity, Significance and Industry involved. There are several ways of collecting the appropriate data which
differs considerably in the context of money, cost, time and other resources at the disposal of the researcher. For the present study, both primary and secondary data have been collected and used for analysis.

Further, in research, there are various points of departure a researcher can choose from. In positivist approach, the researcher assumes that there is a truth to be discovered and that reality is value free, a-historical and cross-cultural. Science should, therefore, be neutral or value free. Further, a careful distinction between scientifically established objective meanings and subjective meanings are made.

The purpose of this chapter is to:

- discuss our research philosophy in relation to other philosophies;
- expound our research strategy, including the research methodologies adopted;
- introduce research instruments that we have developed and utilized in the pursuit of our research.

### 4.2 RESEARCH PHILOSOPHY

A research philosophy is a belief about the way in which data about a phenomenon should be gathered, analyzed and interpreted. The term epistemology (what is known to be true) as opposed to doxology (what is believed to be true) encompasses the various philosophies of research approach. The purpose of science, then, is the process of transforming things believed in into things known: doxa to episteme. Two major research philosophies have been identified in the Western tradition of science, namely,
the positivist (sometimes called scientific) and interpretivist (also known as anti-positivist) (Galliers 1991).

4.3 METHODOLOGICAL ISSUES AND THE RATIONALE

While time constraints did not allow the use of the longitudinal approach, which is often the most preferred in such studies, the methodology, simple as it may sound, provides the opportunity to establish a baseline for a future longitudinal assessment. Mitchell (1989, 1990 and 1991) concluded that research on institutional arrangements for resource management has focused on ex-post studies of specific resource management programmes and projects with an emphasis on descriptive as opposed to predictive approaches. The present study, following this tradition, takes on a descriptive as opposed to predictive approach. Similarly, in the process and outcomes analysis, the case study has been the dominant research design. By concentrating on real world case studies, not only the researcher can test the applicability of some framework and model and tools, but he can assess such methods within a relevant context, which should enhance the replicability of the methodology and the results.

Hence, the study has used a post-positivist approach with less emphasis on aspects of modelling, benefit-cost analysis and analytical (statistical) approaches. This post-positivist approach is not biased towards quantification, but addresses adequately issues of uncertainty, values and socio-historical and behavioural contexts. Mitchell (1989) concluded that phenomenological approaches to resource management have proved useful. He characterised phenomenological approaches as those that concentrated on the understanding of the human/environment interface by focusing on human attitudes, experiences and actions. The choice of a post-positivist
methodology is not to mean that the positivist approaches are not valid. They are very valid. In fact, the post-positivist approaches are valid inasmuch as they contribute to the body of theory, of course.

Sjoberg and Nett (1968) highlighted the strengths and weaknesses of the structured and un-structured interview format. They also concluded that structured interviews provide a means to standardise responses, facilitate the verification of theories and hypotheses, and provide greater reliability than unstructured interviews. They also noted that structured interviews can introduce bias, as researcher may impose their own categories and may have a tendency to oversimplify reality.

4.4 A REVIEW OF RESEARCH METHODOLOGIES

Qualitative Research (Interpretive): Interpretivists contend that only through the subjective interpretation of, and intervention in, reality can that reality be fully understood. The study of phenomena in their natural environment is key to the interpretivist philosophy, together with the acknowledgement that scientists cannot avoid affecting those phenomena they study. They admit that there may be many interpretations of reality, but maintain that these interpretations are in themselves a part of the scientific knowledge they are pursuing. Interpretivism has a tradition that is no less glorious than that of positivism, nor is it shorter.

Quantitative Research (Positivistic, Statistical): Quantitative research methods are research methods dealing with numbers and anything that is measurable. Quantitative research methodology calls for what is known as hard data in the form of numbers. Quantitative research methodology is about the collection of data in their numerical form. So they can be easily
measured or counted. This research methodology is highly preferred by the positivist researchers who want to observe the social reality in terms of quantification and objectivity. By quantitative methods, researchers have come to mean the techniques of randomized experiments, paper and pencil “objective” test, multivariate statistical analysis, sample survey and the like (Cook and Reichardt 1979, Neuman 2000). In social sciences, quantitative research methods express different social phenomena in numbers. In the present case of research, the scholar has to use numerical data which are associated with primary education. In the study, the researcher’s primary aim is to collect primary and secondary data. These data help to discover the past states, present states as well as future trends of the area by using available data. This is an additional advantage of using quantitative data.

Also the quantitative research methods help to generalize the existing social phenomena by testing samples. In this research, a sample survey for gathering primary data by using a questionnaire and interviews have been made using the people of the villages in Mysore district of Karnataka. It has been useful to assess and understand how the men and women of the Mysore villages perceive quality of life in general, and evaluate their feelings about a number of aspects of living and working in rural Mysore district, including their well-being, physical, mental and emotional in particular so that some genuine analysis could be done to look at the quality of life of rural people as well as understand how it has been progressing in Karnataka.

Further the researcher has used SPSS (Statistical Package for Social Sciences) computer software to organize and analyze the data which are collected through the field survey. It is easy to handle quantitative data with SPSS programme and graphing, tabulating, and describing datasets. When
making graphs using numerical data, it can be understood by anyone who
does not have even a rudimentary knowledge of statistics. However, in the
present research, quantitative data are important to explain the nature, value
and knowledge of quality of life among the people of Mysore district and the
conditions living and working there as well as in the state of Karanataka,
using Mysore district as a case in point.

However, there are some limitations the researcher could identify in
quantitative methods. Quantitative methods cannot clearly explain human
feelings and thoughts like, for example, people’s ‘quality of life’ experiences.
It is difficult to convert such things into numerical data. To overcome this
problem, the researcher has used scaling for measuring certain contexts related
items in the questionnaire. It helps respondents to have freedom to select a
scale provided by the researcher in the tool.

Also dealing with a huge quantity of numbers sometimes makes for
mistakes when handling them. Another limitation of quantitative research
methods is that sometimes we cannot get the exact answers. However,
quantitative data are important to our research in many ways and help to make
general evaluation, regarding policy changes in quality of life related
management, for example, impacting upon the local people’s involvement in
their life and work in the study area.

**Positivism:** Positivists believe that reality is stable and can be observed and
described from an objective viewpoint (Levin, 1988); that is, without
interfering with the phenomena being studied. They contend that phenomena
should be isolated and that observations should be repeatable. This often
involves manipulation of reality with variations in only a single independent
variable so as to identify regularities in, and to form relationships between,
some of the constituent elements of the social world.
Predictions can be made on the basis of the previously observed and explained realities and their inter-relationships. **Positivism has a long and rich historical tradition. It is so embedded in our society that knowledge claims not grounded in positivist thought are simply dismissed as unscientific and therefore invalid** (Hirschheim 1985). This view is indirectly supported by Alavi and Carlson (1992) who, in a review of 902 Information Science research articles, have found that all empirical studies are positivist in approach. Positivism has also had a particularly successful association with the physical and natural sciences. Some of the social science researches are no exception.

There has, however, been much debate on the issue of whether or not this **positivist paradigm** is entirely suitable for the social sciences (Hirschheim 1985), many authors calling for a more pluralistic attitude towards research methodologies (see for example, Kuhn 1970, Bjorn-Andersen 1985, Remenyi and Williams 1996). While we would not elaborate on this debate further, it is germane to our study.

Indeed, some of the difficulties experienced in academic research such as the apparent inconsistency of results, may be attributed to the inappropriateness of the positivist paradigm for the domain. Likewise, some variables, or constituent parts of reality, might have been previously thought unmeasurable under the positivist paradigm - and hence went unresearched (after Galliers 1991).

### 4.5 MEASURING QUALITY OF LIFE AND WORK

It is often difficult to measure quality of life. Almost all measurement tools have multipledomains, with multiple items in each domain. A number of measurement methods have been used for assessing quality of life, for example, for persons with disabilities, including surveys
and questionnaires (for example, Cummins et al 1994, Ferrans and Powers 1985), interviews (for example, Park 1985, Lehman 1988), vicarious interviews, and vicarious surveys (for example, Ouellette-Kuntz and McCreary 1996). Most other researchers have put in efforts at involving the persons with disabilities, but they have depended on avicarious response. In some tools, parents or siblings were the major vicarious respondents for the measurement (for example, Becker et al 1993, Ouellette-Kuntz and McCreary 1996).

As individuals are unique, the uniqueness of each individual is at the heart of how quality of life is measured, especially when they are highly diverse as well. At the individual level, a prominent measurement consideration is whether the person has a disability or not. Schalock (2000) has argued that quality of life for persons with disabilities encompasses the same indicators that are important to persons without disabilities. On the other hand, Hatton (1998) has asserted that the experiences of persons with disabilities are restricted because of the limits imposed by disability conditions; and the limited experiences do result in different indicators of quality of life. Hence, specific attention needs to be paid to the uniqueness of each individual, in conceptualizing and constructing a valid measurement for quality of life (Borthwick-Duffy 1996).

4.6 RESEARCH DESIGN OF THE STUDY

The present research has depended on both the primary sources of data and secondary sources of data (Figure 4.1).
Figure 4.1 Flowchart for methodology shows various stages of the research work

**Primary Sources of Data:** The primary sources are the migrants in Chennai and they have been drawn from the ten administrative zones with 155 city wards (now it is 15, with 200 city wards). They are a random sample of migrants, selected using a snowballing process, in which certain number of migrants have been chosen as they are known to the scholar and the rest of them have been chosen using them to create ‘a snowball of a sample’. They are however chosen at random, from those available for interview.

**Sample and Sampling:** A sample is some part of a larger body specially selected to represent the whole. Sampling is the process by which samples for study are chosen. Sampling is taking any portion of a population or universe as a representative of that entire population or universe. For a sample to be
useful, it should reflect the similarities and differences found in the total group. The main objective of drawing a sample is to make inferences about the larger population from the smaller sample. A census is a survey in which information is gathered from or about all members of a population. For the present study, Simple Random Sampling Method (SRS) was used for collection of information from the tourists from selected tourist spots by using the questionnaire.

A sample of 305 migrants has been chosen from every one of the ten zones of the city, comprising of 155 wards. Thus, the sample is widely scattered and represent different parts of the city. The samples chosen have been interviewed using a custom-designed questionnaire, with a distinct number of questions in each of its sections. The samples have been chosen through a snowball sampling procedure, in which the scholar has chosen a select number of known people (20 migrants to the city) living in scattered localities and neighbourhoods and through them the rest of them in such a way they are representatives of the ten city administrative zones and a majority of the city’s 155 wards.

The Questionnaire: The samples chosen have been interviewed using a custom-designed questionnaire, with a distinct number of questions in each of its sections (Appendix 4.1).

Construction of Research Tool: As per the research design, and for the purpose of data collection, the investigator constructed an interview schedule. The prepared interview schedule was subjected to jury opinions. Based on the jury opinion some items were deleted and some others were modified and finally the interview schedule was streamlined.

Gaining Access: One problem is to gain access to respondents, and the problems of being able to study them, and gain some familiarity with their
world, or share their ‘reality’. Social science methodologies propose many different ways in which this can be achieved, and guidance on the extent to which involvement in a respondent’s life world is necessary to for particular descriptive or analytic ends. Practical problems include getting people to speak to the researcher at all, then getting them to be open, co-operative, and sincere in discussing the aspects of their lives that interest the researcher.

More fundamental for research methodology is the reliance on the personal descriptions of the respondents, who may either intentionally conceal or mislead the researcher, or unintentionally mislead them. In studying someone’s life world, it is unlikely that the respondents will be able to comprehensively and thoroughly describe not only their opinions and thoughts, but the details of everyday activities and relationships, and the context in which they conduct them, especially in the space of a relatively short interview.

Participant observation is a method that tries to surmount these obstacles, but at the expense of huge effort by the researcher, and can only be carried out in a situation where the researcher can actually live or work within a small group over an extended period of time. Since the researcher intended to look at a number of groups, and across the social network of one particular member, this type of methodology is impossible. What is more, gaining access to the work place or social clubs may be possible, but living in private homes is very difficult. Only a few researchers have tried this (for example, James Lull in order to examine media use (Lull 1990)).

**Pilot Study:** The items were arranged in a random sampling and administered to a sample of 50 select migrants, chosen from among his neighbours, randomly, in a city ward adjacent to his own. Proper instructions were given before the administration of the questionnaire. This enabled us to identity the vague items, which were ambiguous or difficult to understand. They were
then deleted or refined and rewritten in a way it is understandable to migrants of Chennai.

**Personal details** of the respondents solicited have been from 10 questions whereas the details on migration have been gathered using 8 questions, and details on migrants’ living and working conditions have been collected using 15 questions. The questionnaire has also been designed in a way that there have been three different sections of scaled items, namely, **overall quality of life** with 9 items for scaling, **overall impressions of quality of life** with 5 items of scaling and a longer section of the questionnaire with 32 different items of scaling on eight different aspects of the city: **urban environment, health, education, housing and basic infrastructures, employment, city economy, recreation and safety**.

A typical data matrix represents multiple items or scales (305 cases x 56 variables/items) usually thought to reflect fewer underlying constructs about life and work of migrants in Chennai, some of which in comparison with conditions ‘**before**’ they moved into Chennai.

As indicated above the two sections of the custom designed questionnaire are an evaluation of feelings relative to quality of life (in relation to nine simple items) and feelings relative to overall impressions of life and life and work (5 items) using Likert type of scaling as shown below:

**A. Overall Quality of Life**

**Directions:** Evaluate your feelings relative to the quality of life. Using the 1-7 scale below, indicate your choice of scale with each item by placing the appropriate number alongside that item. Please be open and honest in your responding. (1 = Terrible; 2 = Unhappy; 3 = Mostly dissatisfied; 4 = Mixed; 5 = Mostly satisfied; 6 = Pleased; 7 = Delighted)
1. Your personal life
2. Your wife / husband or ‘significant other’
3. Your co-workers
4. The actual work you do
5. The handling of problems in your life
6. Your ability to adjust to change in your life
7. Your life and work as a whole
8. Overall contentment with your life and work
9. The extent to which your life and work have been as you wanted it

B. Overall Impressions of Well-Being and Life and Work

Directions: Evaluate your feelings relative to the quality of life, using the scale: (5 = Better; 3 = Somewhat better; 2 = Worse)

1. Overall my physical well-being is ______
2. Overall my mental and emotional state is ______
3. Overall my ability to handle stress is_______
4. Overall my enjoyment of life and work is_______
5. Overall my quality of life is_______

This study also draws on data from a questionnaire study of the nature, value and knowledge of eight select groups of conditions now of migrants in Chennai, compared to ‘before’ the migrants’ arrival at the city of Chennai then, (in their past). The scholar has conducted face-to-face interviews with the 305 migrants of varying ages (79.3 per cent of under-30 years, 18.6 per cent of 31-40 years and just about 1.0 per cent of 40 plus
years) and gender (73.4 per cent of men and 26.6 per cent of women). Specifically, the data for the study focus on the following, and of the migrant respondents.

**Conditions compared to ‘Before’ migrants’ coming to Chennai**

**Directions:** Below are certain items for your evaluation (your perceived understanding of the nature, value and knowledge of it). Using the 1-7 scale below, indicate your evaluation of each item by placing the appropriate number alongside that item. Please be open and honest in your responding. (7 = Best; 6 = Better; 5 = Good; 4 = Average; 3 = Bad; 2 = Worse; 1 = Worst)

**A. Urban Environment**

1. Population density / congestion
2. Accessibility to outside world
3. Quality of environment (your actual evaluation of landscapes around you)
4. Social spaces and cleanliness
5. Overall look and feel

**B. Health**

1. Accessibility to services (in the nearest and the specialty services at distances)
2. Affordability of services (within the means of your income, or even insurances)
3. Quality of doctors, nurses, paramedics and others
4. Quality of services rendered (to inpatients and outpatients in general)

5. Distance, time and cost of getting there

C. Education

1. Accessibility to education (in the nearest and the higher services at distances)

2. Affordability of education (within the means of your income, or even savings)

3. Quality of teachers, schools, colleges and other relevant institutions

4. Your experience with educational services (both public and private)

5. Distance, time and cost of getting there

D. Housing and Basic Infrastructure (that goes with it)

1. Quality of your housing and infrastructure

2. Maintenance of the house and infrastructures (promptly attended to, well kept)

3. Affordable, Comfortable and liveable (healthier too)

E. Employment

1. Supports family / household comfortably

2. Challenging and likeable

3. No hazards and sufferings (occupational hazard, struggles to get)
F. Economy (Chennai economy)

1. Satisfies needs and wants, including general luxuries
2. Progressive and forward looking (scope for improvement and increases)
3. Keeps people happy

G. Recreation

1. Opportunities for recreation
2. Use of opportunities (whether or not the individual or household uses such options)
3. Accessibility (getting to recreation)
4. Affordability (cost of it, within the means of individual or household)
5. Time and cost of recreation (actual time and money spent on a regular basis)

H. Safety

1. Individual safety (respondent’s)
2. Women’s and children’s safety
3. Community safety (neighbourhood, societal)

4.7 RELIABILITY AND VALIDITY OF THE TOOL

The reliability of the learning style the interview schedule has been established by calculating the Cronbach Alpha (r=0.75) and the intrinsic validity was established by taking the square root of the reliability Co efficient
that this tool is highly reliable and valid.

**Secondary Data Sources:** The secondary sources of data have generally been census abstracts of various census years, documentary sources of the government agencies, research reports and international and national reports on migration by different organizations such as the UN, International Population Institutes and national institutions of different kinds. Data as well as perspectives have been collated from the reports and documents for the purpose of writing up the thesis.

**Statistical Methods of Analysis:** Among the statistical tools used in the study are (a) the simple frequency and percentage analysis of questionnaire survey data (one-way as well as two-way tables) and (b) the multivariate statistical analysis of factor using principal components approach. In order that the data are amenable to statistical analyses, the questionnaire data have been converted into different datasets (3 datasets) using the MS EXCEL spread sheet and the analyses themselves have been performed using the SPSS package. The two methods are described in some detail below:

**Simple Frequency and Percentage Analysis:** For the purpose of description of sample and respondent related characteristics, a frequency and percentage analysis has been done for all variables extracted from the questionnaire and put into the dataset. First, a simple frequency of each of the fields with column percentages has been made and then two-way tables using certain select pairs of variables have been carried out, in order to measure variations. The analyst begins to explore the data, by measuring the central tendency of the data, and more importantly, the dispersion of the data around this central tendency.
Frequency analysis is particularly useful for describing discrete categories of data having multiple choice or yes-no response formats. This analysis involves constructing a frequency distribution. The only technical requirement of the frequency analysis is that the categories of response be mutually exclusive and exhaustive. This means that the same observation cannot be counted as belonging to more than one response category. The frequency analysis must be exhaustive in the sense that all respondents must fit into a category. The tables so generated are numerous, only select tables are therefore included in the text while others are interpreted so as to show the variations therein.

**Common Factor Analysis:** Factor analysis is a statistical technique designed to analyze the interrelationships within a set of variables by reducing the complex data to an easily interpretable form (Davis, 2002). In multivariate analysis, the bi-variate techniques are extended so that more than two variables can be considered, the ‘m’ variable becoming the ‘m’ axes of the test space. Procedures of multivariate analysis are often concerned with the problem of reducing the original test space to the **minimum number of dimensions** needed to describe the relevant information contained in the original observations. Multivariate procedures differ in the types of original information they preserve. Some understanding of matrix algebra is essential to using and understanding the multivariate analysis.

It is a particular psychometric model that has been in wide use in social sciences. This helps in the study of the logical implications of systematic inter-correlations within sets of tests. However, the social sciences follow just one of the many approaches to the reduction of dimensionality in correlated systems of measurements and the rotation (varimax, a short form for maximizing variance, for example) of a reduced number of axes to more meaningful positions.
The Factor Analysis (FA) is also a classification procedure in that it may be usefully applied to multivariate situations to classifying the N individuals, on the basis of ‘m’ variables. One particular feature of the FA is that ‘p’ underlying factors in the multivariate sample space model is always less than the ‘m’ variables: \( p < m \). The underlying factor dimensions are drawn from the use of inter-correlations system by generating ‘p’ number of scores each for the ‘N’ individuals. The scores may however be drawn from the varimax rotation, which stands for maximizing variance. If we can measure ‘m’ variables with respect to areal units, the scores may be assigned to these areal units for constructing one or more maps showing real areal differences (or regional variations) in respect of ‘p’ reduced dimensions.

The purpose of factor analysis is to interpret the structure within the variance-covariance matrices of the multivariate data collection made on the different indicators related to an eco-city related aspects, including ecofriendliness and sustainability of the city of Gulbarga, in Karnataka, India. The basic mathematical operations in factor analysis may be stated as follows:

\[
Z_j = a_{j1} P_1 + a_{j2} P_2 + \ldots + a_{jm} P_m
\]

where

\[
Z_j = X_j \text{ X mean / } O_j \text{ or standardized variable}
\]

\[
P_i = \text{(i = 1, 2,..m) are the principal components and}
\]

\[
a_j = \text{(j= 1,2,..n) are the coefficients or factor loadings of}
\]

\[
(i = 1,2,..m) ^{jth}
\]

variable relating to the \( i^{th} \) component.

In other words, each factor is nothing but a linear combination of weighted variables which can also be expressed as:

\[
P_i = a_j X_j \text{ where}
\]
Thus, in factor analysis, a data matrix containing measurements on ‘m’ variables for each of ‘N’ observations is analyzed.

The technique uses extraction of the eigenvalues and eigenvectors from the matrices of correlations or covariance. The basic mathematical operations in factor analysis are done with many embellishments on the procedures.

FA is a deep and complex methodology. It is one of the most widely used multivariate procedures. The model is based on several unique assumptions. For one, the precise number of factor is assumed prior to the analysis. The factors extracted, or rather the number of factors, are validated by the variance each of them explain to the total. There is a progressive decline in the value of variances with the increasing number of factor dimensions. The first or the main factor dimension has the highest of the total variance explained and the bipolar the next highest and so on, resulting in progressively declining variance.

The FA requires that ‘p’, the number of factors, be known prior to analysis. This implies that the investigator has some insight into the probable nature of the factors and can predict a suitable number of factors to be extracted.

The eigenvalue operation in factor analysis is performed on a standardized variance-covariance or correlation matrix. Hence, the FA used here is said to be R-mode factor analysis. This assumes not only that all variables are weighted equally, but also allows us to convert the principal component vectors into factors. In larger matrices such as ours, the eigenvalues usually are more uniform for standardized data than for raw data. And to perform the FA, it is necessary that we convert our unit, or normalize
eigenvalue. The result is a factor, a vector, which is weighted proportionally to the amount of total variance it represents.

The elements in the factors are referred to as **factor loadings**. The eigenvalues represent the proportion of the total variance accounted for by the eigenvectors. The factor loadings on the other hand are the correlation values between the old and the new, transformed variables.

If we arrange the factor loadings in a matrix form, we have then a factor matrix. If we square the elements in the factor matrix and sum within each variable, the totals are the amount of variance of each variable retained in the factors. These sums are referred to as the communalities and are symbolically represented as $h_j^2$. The communalities are equal to the original variances.

A specific rule that most factor analysts suggest in the extraction of factor is that of retaining all factors, which have eigenvalues greater than one. That is, retain all factors, which contain greater variance than the original standardized variables. But of course in most instances only a few of the factors will contain most of the variances in the dataset and hence this recommendation is useful. If factor theory is applicable to any given dataset, a few factors should account for a very high percentage of the variance and the communalities of the variables found under each factor dimension is high.

The FA is said to be **reducing the dimensionality** of a problem to a manageable size. However, the meaning of the factors may be difficult to deduce. This problem is overcome by resorting to maximization of the variance of the loadings on the factors. This in other words is maximizing the range of the loadings. This is done in the analysis here by a rotation procedure called Kaiser’s varimax rotation. The rotation of the factor axes is performed, iteratively. The analysis also results in factor scores, which represent
estimates of the contribution of various factors to each original observation (residents). In fact, factors themselves are estimated from these same data. Thus the computation of factor scores is somewhat a circular process and the results are not unique. Factor analysis explains in a sense the interrelationships in a large number of variables by the presence of a few factors (Kaiser 1958, Harman 1960, Lawrence and Upchurch 1983).

The factor extraction is done with a minimum acceptable eigenvalue of >1.0 (Kaiser 1958, Harman 1960). The factor loading matrix is rotated to an orthogonal simple structure, according to varimax rotation. It results in maximization of variance of factor loadings of the variables. This procedure renders a new rotated factor matrix in which each factor is described in terms of only those variables and affords greater ease for interpretation. Factor loading is a measure of the degree of closeness between the variables and the factor. The largest loading, either positive or negative, suggests the meaning of the dimension; positive loading indicates that the contribution of the variables increases with the increasing loadings in a dimension; and negative loading indicates a decrease (Lawrence and Upchurch 1983).

Factor analysis manages over a hundred variables, compensates for random error and invalidity, and disentangles complex interrelationships into their major and distinct regularities. It is mathematically complicated and has diverse and numerous considerations in application. The purpose of this discussion is to enhance the understanding and utilization of the results of factor analysis, rather than provide a technical description.

Conceptually, factor analysis is a means by which the regularity and order in phenomena can be discerned. As phenomena co-occur in space and in time, they are patterned; and as these co-occurring phenomena are however independent of each other, there are a number of distinct patterns
What factor analysis does is this: it takes measurements and qualitative observations and resolves them into distinct patterns of occurrence. It makes explicit and more precise the building of fact-linkages going on continuously in the human mind.

Factor analysis applied to delineate patterns of variation in characteristics is called R-factor analysis. It is applied in order to explore a content area, structure a domain, map unknown concepts, classify or reduce data, illuminate causal nexuses, screen or transform data, define relationships, test hypotheses, formulate theories, control variables, or make inferences (Rummel 1970, Ram 1982, Davis 2002).

When a table of data, say, answers to a questionnaire, are interrelated in a complex fashion, and then factor analysis is used to untangle the linear relationships into their separate patterns. Each pattern appears then as a factor delineating a distinct cluster of inter-related data. It is useful for reducing a mass of information to an economical description. For example, the data are facilitated by reducing them to their common factor patterns. The factors concentrate and index the dispersed information in the original data and can therefore replace the characteristics without much loss of information.

Factor analysis is often employed to discover the basic structure of a domain. It can be used to group interdependent variables into descriptive categories. It is also used to classify, for example, individual profiles into types with similar characteristics or behaviour. Or it is used on data matrices of a social-choice type to show how individuals or social groups cluster on their transactions with or choices of each other.

An investigator often wishes to develop a scale on which individuals, or groups, or even a community can be rated and compared. One
problem in developing a scale is however to weight the characteristics being
combined. The analysis thus offers a solution by dividing the characteristics
into independent sources of variation (factors). Each factor then represents a
scale based on the empirical relationships among the characteristics. The
factor score results are actually such scales, developed by summing
characteristics times the weights (see Distephano et al 2009).

Hypotheses abound regarding dimensions of attitude, personality,
group, social behaviour, and revealed perceptions. Since the meaning usually
associated with ‘dimension’ is that of a cluster or group of highly inter-
correlated characteristics or behaviour, factor analysis is used to test for their
empirical existence. Which characteristics or behaviour is, by theory, related
to which dimensions can be postulated in advance and statistical tests of
significance can be applied to the factor analysis results (Steiger 1990).

Common factor analysis is concerned with defining the patterns of
common variation among a set of variables. Variation unique to a variable is
however ignored. In contrast, another factor model called component factor
analysis is concerned with patterning all the variation in a set of variables,
whether common or unique (Kaiser 1958, Harman 1960, Lawrence and

The Algebraic Model: A traditional approach to expressing relationships is
to establish the mathematical function \( f(X, W, Z) \) connecting one variable,
\( Y \), with the set of variables \( X, W, \) and \( Z \). Such a function might be
\( Y = 2X + 3Z - 2W \), or \( Y = 4XW/Z \). The variables on both the right and the left side of
the equation are known, data are available, and it is only a question of
determining the best function for describing the relationships. Let us assume
that there are several \( Y \) variables and our \( Y \) variables are related to a number
of functions operating linearly. That is, Equation (4.1):
\[ Y_1 = \alpha_{11} F_1 + \alpha_{12} F_2 + \ldots + \alpha_{1m} F_m, \]
\[ Y_2 = \alpha_{21} F_1 + \alpha_{22} F_2 + \ldots + \alpha_{2m} F_m, \]
\[ Y_3 = \alpha_{31} F_1 + \alpha_{32} F_2 + \ldots + \alpha_{3m} F_m, \]
\[ \ldots \]
\[ \ldots \]
\[ \ldots \]
\[ Y_n = \alpha_{n1} F_1 + \alpha_{n2} F_2 + \ldots + \alpha_{nm} F_m, \quad (4.1) \]

where:

\[ Y = \text{a variable with known data} \]
\[ \alpha = \text{a constant} \]
\[ F = \text{a function, } f(\text{)} \text{ of some unknown variables.} \]

This is important in understanding factor analysis to remember that \( F \) stands for a function of variables and not a variable. For example, the functions might be \( F_1 = XW + 2Z \), and \( F_2 = 3X^2Z/W^{1/2} \). The unknown variables entering into each function, \( F \), of Equation (4.1) are related in unknown ways, although the equations relating the functions themselves are linear.

Within the algebraic perspective, what does factor analysis do? By application to the known data on the \( Y \) variables, factor analysis defines the unknown \( F \) functions. The loadings emerging from a factor analysis are the \( a \) constants. The factors are the \( F \) functions. The size of each loading for each factor measures how much that specific function is related to \( Y \). For any of the \( Y \) variables of Equation (4.1) we may write:

\[ Y = \alpha_1 F_1 + \alpha_2 F_2 + \alpha_3 F_3 + \ldots + \alpha_m F_m, \quad (4.2) \]

with the \( F \)'s representing factors and the \( \alpha \)'s representing loadings.
It is possible that some of the F functions are common to several variables. These are called group factors and their delineation is often the goal of factor analysis. Besides determining the loadings, factor analysis also generates data (scores) for each case (individual, migrant in our case) on each of the F functions uncovered. The derived values for each case are called factor scores. They, along with the data on Y and Equation (4.1) give a mathematical relationship among data as useful and important as the classical equations like \( Y = 2X + 3Z \).

**Graphical and Other Representations (SPSS, MSEXCEL):** Graphical representations and charts have been created to illustrate the questionnaire survey data analyzed. Maps used in the thesis have generally been prepared using the modern geographical information systems.

**Library Research:** In addition, library research has been gone through meticulously using different University libraries and also other research institutions such as the Madras Institute of Development Studies and Anna Centenary Library.

**4.8 CONCLUSION**

This chapter has dealt with the research methodology adopted in the study of Characteristics of Migrants in Chennai City, particularly for a sample chosen from 155 city wards organized into 10 city administrative zones. The discussion in the chapter has focused on the qualitative and quantitative research methodologies that could be useful in a study of migrants’ quality of life and then narrowing the focus on the five areas of a methodology, namely, the primary sources of data, the secondary sources of data, the statistical techniques used in the study to bring out the quality of life and work aspects of migrants in Chennai city, using simple frequency and percentage analysis and multivariate common factor analysis, graphical
representations using the SPSS and MS EXCEL software and library research involved in the study. The methods chosen for the current research are most appropriate to assemble ideas, illustrate data and analyse, interpret and infer conclusions from the analysis of questionnaire survey data. In the next chapter, the discussion turns to essentially the socio-demographic and economic details of the migrants to Chennai city and also focuses on the overall quality of life and overall impressions on well-being, life and work and quality of life of the migrants studied with a purpose of elaborately documenting their characteristics for an appreciation of their life and work and their revealed perceptions on some select aspects of urban life.