CHAPTER III

RESEARCH METHODOLOGY

3.1 RESEARCH DESIGN

The research process involves identifying a management problem or opportunity; translating that problem or opportunity into a research problem; and collecting, analyzing, and reporting the information specified in the research problem (Kervin, 2004). A research design is the detailed blueprint used to guide a research study toward its objectives. The process of designing a research study involves many interrelated decisions (Minocha, 2006). The most significant decision is the choice of research approach, because it determines how the information will be obtained. Tactical research decisions are made once the research approach has been chosen (Blumberg and Schindler, 2008). Here the focus is on the specific measurements to be made or questions to be asked, the structure and length of the questionnaire, and the procedure for choosing a sample to be interviewed. These tactical decisions also are constrained by time and budget availability, so before a study can be implemented, estimated costs must be compared to the anticipated value.

One of the best ways to get answers for the research questions set is a well planned, structured research design (Fink, 2003). In this thesis the researcher ensures that the thesis is well structured and well planned right from setting objectives, research hypothesis, framing necessary and support theories for the research objectives and finally a questionnaire is designed with a intent to obtain answers for the set research objectives. According to Bryman and Bell (2007), right from writing hypothesis to final analysis of data including any operational activities a well structured research design is required. An effective research design constitutes the
blueprint for the collection, measurement and analysis of data and ensures that the research is conducted within the conceptual structure.

The research design adopted by the researcher in this thesis is descriptive research design. Descriptive research, also known as statistical research, describes data and characteristics about the population or phenomenon being studied. The data description is factual, accurate and systematic; the research cannot describe what caused a situation. According to Burns (2008), in descriptive research the research hypotheses often will exist, but they may be tentative and speculative. As in this thesis also the researcher has designed hypothesis which has been presented at the end of this chapter and this will be tested in this study in latter part.

In general, the relationships studied will not be causal in nature. However, they may still have utility in prediction. In social sciences; descriptive research usually takes one of two forms: Survey research and Observational research. According to Kervin (2004), “Survey research is the systematic gathering of information from respondents for the purpose of understanding and predicting some aspect of the behavior of the population of interest. As the term is typically used, it implies that the information has been gathered with some version of a questionnaire. The survey researcher must be concerned with sampling, questionnaire design, questionnaire administration, and data analysis”. And accordingly based on Kervin’s quote on survey, the researcher is careful enough to decide about questionnaire design and the sampling technique which is discussed later in this chapter.
3.2 DATA COLLECTION

3.2.1 Types of Data Collected

The researcher has used both primary and secondary data for his research. Secondary data means data that are already available i.e., they refer to the data which have already been collected and analyzed by someone else (Chisnall, 2007). When the researcher utilizes secondary data, then he has to look into various sources from where he can obtain them. In this case he is certainly not confronted with the problems that are usually associated with the collection of original data. Secondary data may either be published data or unpublished data (Carson and Perry, 2001). Usually published data are available in various publications of the central, state are local governments; various publications of foreign governments or of international bodies and their subsidiary organization; technical and trade journals; books magazines and newspapers; reports and publications of various associations connected with business and industry, banks, stock exchanges; reports prepared by research scholars, universities, economists, in different fields; and public records and statistics, historical documents, and other sources of published information (Downie, 2003).

The primary data are those, which are collected for the first time and thus happened to be original in character. Such data are collected with specific set of objectives to assess the current status of any problem. Primary data collection is necessary when a researcher cannot find the data needed in secondary sources (Cassel and Johnson, 2006). Market researchers are interested in primary data about demographic, socioeconomic characteristics, attitudes, opinions, interests, awareness/knowledge, intentions, motivation, and behavior. Primary data are
originated by the researcher for the specific purpose of addressing the problem at hand (Collis and Hussey, 2009). Since primary data is collected with specific purpose, it forms the most significant data of the entire thesis and it is ultimately used for the purpose of analysis.

3.2.2 Methods of data collection

3.2.2.1 Collection of secondary data

The secondary data is collected from the websites, international and national journals in the field of management as well as marketing, business magazines, business dailies, referred text books in marketing management as well as service quality and academic studies conducted in the related areas for the purpose of building a strong conceptual background including the review of literature for the study.

3.2.2.2 Collection of primary data

There are several methods of collecting primary data, particularly in surveys and descriptive researches. Some of the methods are observation method, interview method, questionnaire method, scheduling method, and other methods like warranty cards, distributor audits, pantry audits, consumer panels, mechanical devices and so on. For the descriptive type of researches, the best – suited research approach for collecting primary data is the survey technique using questionnaire method. From a sample, data is collected and the different magnitudes are measured with respect to the whole population (Cooper, 2006). Questionnaire method of data collection is quite popular, particularly in case of big enquiries. It is being adopted by private individuals, research workers, private and public organization and even by governments (David Aaker, Kumar and George Day, 2000). In this thesis the
researcher has used questionnaire method for the purpose of collecting primary data from the retail customers of both the organized and unorganized sector to know the service quality gap with a view to provide a service quality module to have a better performance so that the benefits of the growing retail industry can be tapped to the fullest extent by the retailers.

3.2.3 Questionnaire Design

As discussed earlier for the purpose of collection of primary data questionnaire method has been adopted in this thesis. But designing and constructing a questionnaire for effective implementation is not an easy job for the researcher (Peterson, 2000)\textsuperscript{15}. It is one of the most interest and challenging tasks of conducting a research for the researcher. Questionnaire is a self administered measuring instrument although it is designed by the researcher, for obtaining data from the respondents with view to find out their attitude, opinion to answer the research objectives set earlier in the study. Questionnaires are not just used to record the responses of the respondents regarding their attitude and opinion but also are used to describe, compare, or explain individual and societal knowledge, feelings, values, preferences and behavior (Bryman and Bell, 2007)\textsuperscript{16}. Since this thesis has adopted descriptive research design, participation of large number of respondents are essential to come to a conclusion at the end of the study so the viable method will be the questionnaire method. This is because, the questionnaire method of communication allows participation of a larger number of respondents to be targeted as questionnaires are distributed more widely to gain a broad viewpoint of the study.
This thesis has used both attitudinal and classification questions. Classification questions are directed to the participants of the survey with intent to gather demographic and socio-economic related things (Donald Tull and Dell Hawkins, 2005). The data collected through this classification consider age, gender, social class, occupation, family size and so on. To obtain basic information concerning the background of respondents this kind of classification questions are designed. This will be essential when the data is presented for analysis, as this type of information will consist of the respondents personal details such details are crucial in understanding, why the respondents may react to some questions when it comes to providing opinion on the shopping experience, in either an organized retailing or unorganized retailing to analyze the service quality of the service provider that is the retailer. Attitudinal questions are designed with intent to seek opinions or basic beliefs which people have about the shopping experience in the organized as well as the unorganized retail industry to analyze the quality of the performance of the service providers in the retail industry with a view to know the service quality gap and suggest a service quality module. It is highly advantageous to apply this question type where information about people’s preferences and views about the service quality performance from the unorganized and the organized retailers. Use of attitude questions allows the researcher to compare and contrast many different opinions and beliefs and assess how those attitudes and beliefs are conceived. (Bryman and Bell, 2007).

As Classification and attitudinal questions are two types of technique adopted by the researcher while designing a questionnaire, similarly there are two other types for developing a questionnaire, one is the open ended questions and the other one is the close ended questions. The idea of open and closed questions is to allow the researcher to determine how best to guide the way in which respondent’s answers the
questions framed. By framing an open ended questionnaire, it would give the participant an opportunity to express in detail about the questioned posed to the respondent. In other words, respondents are free to reply in their own words rather than being limited to choosing from among a set of alternatives. This definition implies that respondents have full control of their opinion, resulting in a clear advantage for the researcher, as an unbiased research response is gained. But for a thesis which required conclusion through analysis, this is not a suggested technique, rather closed questions can be used by the researcher. In the close ended questionnaire, there are a number of options for the participants to the questions asked by the researcher from which the respondents are to select one or more (Hussey, 2007)\(^{19}\). This has advantages in the analysis stage as it would provide an opening in which the respondent would be able to express their opinion, allowing for a stronger correlation to be gained (King, 2004)\(^{20}\). In this thesis the researcher has designed a close ended questionnaire for collecting data from the customers from both the organized and unorganized retailers in Chennai with the objective of knowing their demographic profile and identifying how they perceive the quality of the service provided by the organized and the unorganized retailers in the retail industry.

The questionnaire also has two sections; section 1 is comprised of questions that require some measure of personal details in relation to the age, gender, job or position in the organization. Section 2 is made up of likert scale questions. A likert scale is defined by Zikmund (2003)\(^{21}\) as “A measure of attitudes designed to allow respondents to indicate how strongly they agree or disagree with carefully constructed statements that range from very positive to very negative towards an attitudinal object”. They require the respondents to give their perspective on the various factors relating to the shopping experience from the customers from both the organized and
unorganized retailers in Chennai, based on the respondent’s agreement level, satisfaction level and the motivation level. (Refer Appendix – survey questionnaire).

### 3.2.4 Cronbach's alpha test

The researcher has used Cronbach's alpha test as a measure of internal consistency, that is to check the questionnaire, as to how closely related a set of items are as a group. A "high" value of alpha is often used (along with substantive arguments and possibly other statistical measures) as evidence that the items measure an underlying (or latent) construct. However, a high alpha does not imply that the measure is one-dimensional (King, 2004). If, in addition to measuring internal consistency, you wish to provide evidence that the scale in question is one-dimensional, additional analyses can be performed. Exploratory factor analysis is one method of checking dimensionality. Technically speaking, Cronbach's alpha is not a statistical test and it is a coefficient of reliability (or consistency). This test is used to test the consistency of the attitude questions from the respondents to check the reliability of the questionnaire.

### 3.3 SAMPLING PROCESS

According to Kervin (2004), “Sampling is a necessary and inescapable part of human affairs. If all possible information needed to solve a problem could be collected, there would be no need to sample. Any researcher can rarely do this, however, because of limitation on the amount one can afford to spend and on the available time, or for other reasons. Therefore, the researcher must take samples”. Sampling design includes the sampling unit, sample population, sample size and the sampling method employed for identifying the potential respondents.
3.3.1 Sampling Unit and Choice of Study

The Sampling Unit is the customers from organized retailing and customers from unorganized retailing in Chennai, India.

Chennai is the fourth largest city in India by area and the Capital city of Tamilnadu. The City’s focus on the manufacturing sector has led to strong infrastructure development on the city. Good intra city connectivity and presence of an efficient air, road, rail and sea networks across regions have added to the attractiveness of the city. The city also has a high urbanization rate reflecting equitable development and growth in the region.

According to India Organized retail Market Diagnosis and Outlook 2011, Chennai is ranked top 8 in favorite destinations for retailers in India among the Southern region. Organized retail in the city is developed at a much slower rate as compared to other metros of the country. This can be primarily attributed to the traditional nature of the people here and the lack of encouragement shown towards new retail ventures by them. This attitude is reflected in the growth of retail in the city where the high street format (unorganized retail) dominates the retail market scenario. The reason for the lack of organized retail development is because of the consumer buying pattern which is more value driven also with the domination of Unorganized retail. The retail development in Chennai which had stagnated over the past couple of years is expected to witness steady growth in the future. This has stimulated to choose, Chennai as the choice of the study.
Table No. 3.1  
Favorite Destinations for Retailers in India

<table>
<thead>
<tr>
<th>S.No</th>
<th>City</th>
<th>Region</th>
<th>% of Retail Outlet Center</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mumbai</td>
<td>W</td>
<td>7.0</td>
</tr>
<tr>
<td>2</td>
<td>Bangalore</td>
<td>S</td>
<td>6.5</td>
</tr>
<tr>
<td>3</td>
<td>Kolkata</td>
<td>E</td>
<td>5.9</td>
</tr>
<tr>
<td>4</td>
<td>Delhi</td>
<td>N</td>
<td>5.7</td>
</tr>
<tr>
<td>5</td>
<td>Hyderabad</td>
<td>S</td>
<td>5.7</td>
</tr>
<tr>
<td>6</td>
<td>Ahmadabad</td>
<td>W</td>
<td>5.4</td>
</tr>
<tr>
<td>7</td>
<td>Pune</td>
<td>W</td>
<td>5.3</td>
</tr>
<tr>
<td>8</td>
<td>Chennai</td>
<td>S</td>
<td>5.0</td>
</tr>
<tr>
<td>9</td>
<td>Noida</td>
<td>N</td>
<td>4.2</td>
</tr>
<tr>
<td>10</td>
<td>Gurgoan</td>
<td>N</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Source: Dun & Bradstreet India, 2011

By the report of India Organized retail Market Diagnosis and Outlook 2011, Zone locations in Chennai has been classified into Central Business District (CBD) and Off- Central Business District (Off-CBD) based on the number of retailers both organized and unorganized as well as the density of retail customers. This helped the researcher to choose the sample area for the data collection. The Survey is conducted in the areas of Thiyagaraya Nagar, Nungambakkam, Mylapore, Alwarpet and Egmore which is termed as Central Business District in Chennai.

3.3.2 Sampling Population

In this thesis the city of Chennai is considered for the purpose of study, the total population according the census data of 2011 shows that the population of Chennai is at 46,81,086 and it is an established fact that the entire population will do some sort of retailing .
3.3.3 Sample size

Since the population size in Chennai is more than 50,000 the researcher has adopted the following sample.

Sample Size Formula

\[
SS = \frac{Z^2 \times P \times (1 - P)}{C^2}
\]

\(Z\) = \(Z\) value (e.g. 1.96 for 95% confidence level)

\(P\) = percentage picking a choice, expressed as decimal (0.5 used for sample size needed)

\(C\) = confidence interval, expressed as decimal (e.g. 0.04 = ± 4)

Correction for finite population:

\[
\text{New } SS = \frac{SS}{1 + \frac{SS - 1}{POP}}
\]

\(POP\) = population

Since the population is finite as the population is more than 50,000 the researcher has not considered the correction factor, applying the formula of SS by substituting 1.96 for \(Z\), 0.5 for \(P\) and .04 for \(C\) the sample size arrived are 600 which is the sample size for this thesis. Since the researcher has adopted a comparative study between the customers from organized retailing and unorganized retailing, to have an equal representation from these two sub groups, 300 customers of organized retailing and 300 customers of unorganized retailing is considered for this study.

3.3.4 Sampling Technique

Sampling is the use of a subset of the population to represent the whole population (Kothari, 2005). Probability sampling or random sampling is a sampling technique in which the probability of getting any particular sample may be calculated. Non probability sampling does not meet this criterion and should be used
with caution. Non probability sampling techniques cannot be used to infer from the sample to the general population. The difference between non probability and probability sampling is that non probability sampling does not involve random selection and probability sampling does. Does that mean that non probability samples aren't representative of the population? Not necessarily. Non-probability sampling is a sampling technique where the samples are gathered in a process that does not give all the individuals in the population equal chances of being selected. But it is applied because in any form of research, true random sampling is always difficult to achieve. Most researchers are bounded by time, money and workforce and because of these limitations, it is almost impossible to randomly sample the entire population and it is often necessary to employ another sampling technique, the non-probability sampling technique. It can be used when randomization is impossible like when the population is almost limitless. The advantage of non probability sampling is its lower cost compared to probability sampling. Many research practice many analysts draw generalizations (e.g., propose new theory, propose policy) from analyses of non probability sampled data. So being a widely accepted phenomenon the researcher in this thesis has adopted non probability sampling.

The Sampling Method that was chosen in this thesis is entirely non-probabilistic in nature. The main reason behind this the entire and exact population size is unknown for the researcher and due to the difficulties in adopting probability sampling in the city of Chennai finding out the consumers in the retail industry, thereby the non-probability type of sampling method is adopted in this study. Under the non-probability sampling the researcher has adopted convenience sampling which also known as availability is sampling. Availability sampling is where the researcher
selects subjects on the basis of availability. Convenience sampling is also called haphazard sampling which includes interviewing people who based on their availability in a haphazard fashion since no probability or some kind of system is followed. This is also known as convenience sampling. If a quick reaction to a preliminary service concept is desired to determine if it is worthwhile to develop it further, a convenience sample may be appropriate. (Zikmund, 2003)\textsuperscript{25}. According to Fink Arlene (2003)\textsuperscript{26}, “A convenience sample is one in which the only criterion for selecting the sampling units is the convenience of the sampler”. The researcher for the purpose for the study has used convenience sampling in this thesis to obtain information quickly and inexpensively.

3.4 QUESTIONNAIRE PRE-TESTING

Kervin (2004)\textsuperscript{27} articulate that the reliability of the questionnaire is concerned with the consistency of responses to the questions. The internal consistency method will be used to measure the consistency of responses across either all questions or a subgroup of questions from the questionnaire. Prior to the distribution of the questionnaire, it was tested on selected sample; it is important to embark on this exercise to make sure that no questions or aspects of the questionnaire posses any problem or confuses the respondent. This forms an important part of the research methodology stage. Certain faults in the content, structure and design of the questions that were not recognized when designing the questionnaire are dealt at this stage and this can make the questionnaire more potent in delivering the required results. There are three main things that should be carefully checked; flow, timing and respondent interest and attention. (Fink Arlene, 2003)\textsuperscript{28}. In order to test the questionnaire design for collecting data, researcher had earlier issued for pre-test purpose ten questionnaires to the customers from retail industry. There were no issues here as all
respondents that took part in the test understood the questions, and said they could very well understand the language.

Another element to test was the timing. Examining the questionnaire on face value, some of the respondents said that the questionnaire looked too long, even though it was not time consuming and easy to answer as many of the respondents later testify. Also the interest and attention of the respondent was another crucial area to be tested. To keep the appeal alive, the researcher made sure that enough time was given to every respondent and attended a few sessions with them, in some cases on individual basis to make the respondent appreciate the significance of the research. After testing the questionnaire on a small sample to ensure it has been drafted in the most constructive manner, this questionnaire was then distributed to the selected sample.

3.5 **TOOLS USED FOR ANALYSIS**

Descriptive statistics are used to describe the basic features of the data in this study. They provide simple summaries about the sample and the measures. Together with simple graphics analysis, they form the basis of virtually every quantitative analysis of data in this study. Descriptive Statistics is used to present quantitative descriptions in a manageable form in this study. And following are the statistical tools used for the purpose of data analysis.

3.5.1 **Percentage Analysis**

According to Walker (2007), the percentage analysis is the method to represent raw streams of data as a percentage (a part in 100 - percent) for better understanding of collected data.
Percentage can also be used to compare the relative term, the distribution of two or more series of data. In the major part of the analysis chapter, percentage Analysis has been used.

**Formula**

\[
\frac{X \times 100}{N}
\]

In the thesis the sample chosen is 600. So the data collected after the survey will be converted into the percentage.

### 3.5.2 Analysis of variance

According to Burns (2008), ANOVA is a particular form of statistical hypothesis testing heavily used in the analysis of experimental data. A statistical hypothesis test is a method of making decisions using data.

A test result (calculated from the null hypothesis and the sample) is called statistically significant if it is unlikely to have occurred by chance alone. A statistically significant result (when a probability (p-value) is less than a threshold (significance level)) justifies the rejection of the null hypothesis. In the typical application of ANOVA, the null hypothesis is that all groups are simply random samples of the same population. This implies that all treatments have the same effect (perhaps none). Rejecting the null hypothesis implies that different treatments results in altered effects. By construction, hypothesis testing limits the rate of Type I errors (false positives leading to false scientific claims) to a significance level.

**Formula**

\[
F = \frac{nS^2}{S^2}
\]
In this thesis the researcher also has used ANOVA both one way and two way for the purpose of testing of the hypothesis that the variation in an experiment is no greater than that due to normal variation of individuals' characteristics and error in their measurement. The reason for doing an ANOVA is to see if there is any difference between groups on some variable.

In this thesis ANOVA is used to test the following research hypothesis.

- Relationship between occupation of the respondent and most frequently shopped item.
- Relationship between income of the respondent and most frequently shopped item.
- Relationship between age of the respondent and most frequently shopped item.

3.5.3 Chi Square Analysis

According to Cooper (2006), chi square is a non-parametric test of statistical significance for bivariate tabular analysis. To test the hypothesis Chi-squared test is used as a statistical tool in this project. Any appropriately performed test of statistical significance lets the researcher to understand the degree of confidence that can have in accepting or rejecting a hypothesis. Also it invokes no assumption about the form of original distribution from which the observations are made. In this method we test if two attributes considered are dependent or not.

Typically, the hypothesis tested with chi square is whether or not two different samples that is different enough in some characteristic or aspect of their behavior that the researcher can generalize from the samples that the populations from which the samples are drawn are also different in the behavior or characteristic. Also it invokes no assumption about the form of original distribution from which the observations are made.

109
**Formula**

\[ \chi^2 = \sum \frac{(o - e)^2}{e} \]

Chi-square test is used as a statistical tool in this thesis. In this method test is conducted to find out whether the two attributes or the variables considered are dependent or not.

In this thesis chi square analysis is used to test the following research hypothesis.

- Relationship between occupation, income and gender of the respondents and choice of retailing
- Relationship between age, gender, occupation and income of the respondents and type of stores chosen in organized retailing.
- Relationship between age, gender, occupation and income of the respondents and type of stores chosen in unorganized retailing.

### 3.5.4 Correlation Analysis

According to Kervin (2004)\textsuperscript{32}, correlation, also called correlation coefficient indicates the strength and direction of a linear relationship between two random variables. In general statistical usage, correlation or co-relation refers to the departure of two variables from independence, although correlation does not imply causation. In this broad sense there are several coefficients, measuring the degree of correlation, adapted to the nature of data. According to probability theory and statistics, a number of different coefficients are used for different situations.

Correlation analysis has been used to test the relationship of selected variables. This is done to indicate the strength and direction of a linear relationship between two random variables. In general statistical usage, correlation or co-relation refers to the departure of two variables from independence, although correlation does not imply
causation. In this broad sense there are several coefficients, measuring the degree of correlation, adapted to the nature of data. A number of different coefficients are used for different situations. The best coefficient technique is the Karl Pearson technique which is obtained by dividing the covariance of the two variables by the product of their standard deviations.

**Formula**

\[
r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n \sum x^2 - (\sum x)^2} \sqrt{n \sum y^2 - (\sum y)^2}}
\]

In this thesis correlation is used to find out the extent of relationship between the following variables.

- Age, income, occupation and gender of the respondent and rating of the overall shopping experience in the retail store (organized).
- Age, income, occupation and gender of the respondent and rating of the overall shopping experience in the retail store (unorganized).

### 3.5.5 Factor Analysis

In this thesis the researcher also has used factor analysis for analyzing the opinion of the respondents. Factor analysis describes the variability among the observed correlated variables in terms of a potentially lower number of unobserved variables called factors. In other words, it is possible, for example, that variations in three or four observed variables mainly reflect the variations in fewer such unobserved variables. Factor analysis searches for such joint variations in response to unobserved latent variables. The observed variables are modeled as linear combinations of the potential factors, plus err terms. The information gained about the interdependencies between observed variables can be used later to reduce the set of variables in a dataset.
3.6 SUMMARY OF THE CHAPTER

This chapter focused on all the areas which have to be implemented in order to produce an effective research methodology right from choosing the research design, deciding the type of data to be collected, hypothesis to be tested and finally the type of statistical tools used to analyze the data collected through the questionnaire method. Descriptive research design is the chosen research design for this thesis. Primary and secondary data has been used. Questionnaire is the chosen data collection method, to obtain quantitative data. Both classification and attitudinal questions are used to collect the data from the respondents using closed questionnaire type. The sampling method chosen for this study is stratified random sampling technique which is probabilistic in nature, whereby every member in the population are represented through the sampling process. The next chapter will discuss the results obtained on the field and analyze the results using tables, charts and graphs. The chapter will provide the basis for this paper as a primary and quantitative research.


