CHAPTER 4
RESEARCH DESIGN & METHODOLOGY

4.1 Introduction
In this chapter the philosophical opinion of the researcher is explored which will clarify the reasons for the choice of methodology used in this research. Therefore, the main purpose of this chapter is to present the research methodology and methods used in this study. To answer the research questions the existing most accepted scale for measurement of retail service Quality (RSQS) by Dabholkar, Thrope and Rentz, (1998) was tested for its reliability and validity which was found negative. Then a detailed TOM survey and expert opinion were collected to get the extended instrument for developing the scale for measuring Grocery Retail Store Service Quality in the main study and this new scale was used for the analysis of different relations with overall service Quality. The process used to modify the research instrument and the techniques used to collect and analyze the data will be discussed in detail.

Research can be classified in terms of their purpose. Accordingly, they are most often classified as exploratory, descriptive or explanatory (Sauders, Lewis & Thornhill 2007). Exploratory research is useful when the research questions are vague or when there is little theory available to guide predictions. At times, researcher may find it impossible to formulate a basic statement of the research problem. Exploratory research is used to develop a better understanding (Hair, Babin, Money & Samouel 2003). Exploratory studies are a valuable means of finding out what is happening, to seek new insight, to ask questions and to assess phenomena in a new light. It is particularly useful if researcher wish to clarify the understanding of a problem. There are three principle ways of conducting exploratory research: a search of the literature, talking to experts in the subject, conducting focus group interviews (Saunders, Lewis & Thornhill 2003). Descriptive Research describes some situation. Generally things are described by providing measures of an event or activity. Descriptive Research designs are usually structured and specifically designed to measure the characteristics described in a research question. Hypotheses, derived from the theory, usually serve to guide the process and provide a list of what needs to be measured (Hair, Babin,
Money & Samouel 2003). The object of descriptive research is to portray an accurate profile of persons, events, or situations. It is necessary to have a clear picture of the phenomena on which researcher wish to collect data prior to the collection of the data (Saunders, Lewis & Thronhill 2003).

Studies that establish causal relationships between variables may be termed explanatory studies. The emphasis here is on studying a situation or a problem in order to explain the relationship between variables (Saunders, Lewis & Thronhill 2003). Explanatory studies are designed to test whether one event causes another (Hair, Babin, Money & Samouel 2003).

The purpose of the research is mainly descriptive and explanatory. It is descriptive because descriptive data has been collected through detailed questionnaire and it is also explanatory since the researcher will explain the relationship between service quality and demographic variables. It is somewhat exploratory nature since we are exploring the relationship between service quality variables and satisfaction based on the previous theory to develop a better understanding about the research area.

The chapter begins with the elements of the research process which include the research assumption, theoretical perspective, the methodology and methods used in the research. Subsequently, the chapter explains the methods used for data collection and analysis. The steps for RSQS scale testing include a survey in Delhi, NCR region. Then an extensive survey for TOM factors of the respondents was done. After that the expert opinion were collected to produce an extended instrument to develop the modified scale for the main study.

4.2 Elements of the Research Process

Research design is the research process that involves the overall assumptions of the research to the method of data collection and analysis (Creswell, 2009). The choice of research design depends on the objectives of the research in order to be able to answer the research questions (Crotty, 1998). Therefore, the researcher describes the research elements in general before describing his own philosophical stance. The four elements of the research process, Research Assumption, Theoretical perspective, Methodology and Methods are presented in Table 4.1.
Table 4.1: Elements of the Research Process

RESEARCH ASSUMPTION
The assumption of knowledge that is part of methodology (Crotty, 1998; Creswell, 2009).

THEORETICAL PERSPECTIVE
The researcher’s Philosophical OPINION that is the basis for his/her methodology (Crotty, 1998, Creswell, 2009).

METHODOLOGY
The reason for choosing a method and how the method is used (Crotty, 1998; Collis & Hussey, 2003).

METHODS
Tools and techniques used to gather and analyze data (Crotty, 1998; Saunders, Lewis & Thornhill, 2009).

4.2.1 Overview of the Researcher’s research process
After describing the four elements of the research process the researcher gives an overview of his own research process. The first part (Table 4.1) shows that the researcher of this study follows an objective research assumption, a positivist approach and used surveys for data collection. The second part (Table 4.2) shows a more detailed description of the steps that led to the new modified scale and finally the discussions and findings.
4.3 Research Assumptions and Theoretical perspective

Research assumption provides the opinion upon which the researcher builds the methodology that is used in his/her research (Crotty, 1998; Creswell, 2009). Therefore, it is important that the researcher explains the philosophical position that is being adopted in the research (Crotty, 1998). This could be fundamentally different from one researcher to another (Blaikie, 2007). The researcher's beliefs and understanding of the world influences the research design (Gill & Johnson, 2005). Moreover, research assumption concludes that there is a degree of certain knowledge that can be uncovered (Hanson & Grimmer, 2007).

Research assumption can be shown as a continuum with objectivism at one end and subjectivism at the other (Johnson & Duberley, 2000). Objectivism assumes that the researcher and the area being researched are independent. This means that the researcher has the capability to study the
area of research independently without being influenced by his own views (Johnson & Duberley, 2000). Researchers that follow an objective assumption are mainly concerned with a positivist approach which was historically used for natural sciences such as biology and physics and later adopted by social science (Crotty, 1998). Subjectivism assumes that the researcher and the area being researched are not independent and that the researcher might be involved in a participative enquiry (Collis & Hussey, 2003).

Many philosophers believe that positivism is the foundation and rationale for most management research these days (Johnson & Duberley, 2000). Many social science approaches today have drawn their ideas from positivism or the contradiction to positivism despite the fact that it has been challenged (Smith, 1998; Johnson & Duberley, 2000; Gill & Johnson, 2005). For this reason, some aspects of positivism have become accepted ways of judging truth and reality and are now part of Western culture (Johnson & Duberley, 2000). Consequently, positivism has influenced many approaches that are thought to be an objective way of viewing reality, because of this wide influence, the definition of positivism has become unclear (Smith, 1998).

There are eight main features of positivism (Hussey & Hussey, 1997): First, a quantitative method is usually used, although it is possible to use qualitative methods. Second, large samples are usually used in order to generalise the findings. Third, hypotheses are tested; a hypothesis is usually tested by statistical analysis and a decision is made to reject or accept the null hypothesis. Fourth, data is specific; researchers attempt to measure the data and reach conclusions. Fifth, location; investigations of the problem are not usually conducted in the field. Sixth, reliability; concerned with the repetition of the test, if the test is repeated will the same results be obtained? Seventh, validity; do the findings represent the real situation? Eighth, generalizations are made from the samples; here confidence tests may be used to suggest that the researcher is 95% - 99% confident that the sample represents the population.

In contrast, the interpretive approach has a subjective epistemology that usually follows inductive logic and uses qualitative methods reliant on investigating theories. The theory may be an
existing one or one that the researcher creates in order to have in-depth knowledge of the main study. Reliability is low and validity is high in the findings of the interpretive approach because in the interpretive approach reliability is not as important as it is in the positivist approach (Collis & Hussey, 2003). In the interpretive approach researchers are subjective in the way they look at their findings and attempt to understand and describe the way people view the world (Creswell, 2009). Moreover, the researcher in the interpretive approach understands that his interpretation of these findings is influenced by his own culture, beliefs and experiences (Creswell, 2009).

Deductive logic places emphasis on —arguing from the general to the particular area of research (Plano Clark & Creswell, 2008). The researcher forms a theory based on a hypothesis allowing the researcher to deduce conclusions, which are then examined to establish whether they add to the general knowledge (Blaikie, 2007). The hypothesis that is formed must be tested empirically and the researcher must identify the method of data collection (Bryman & Bell, 2007). According to Bryman & Bell (2007) the deduction process is as follow: theory, hypothesis, data collection, findings, hypothesis confirmation or rejection and revision of theory.

On the other hand, inductive logic places emphasis on —arguing from the particular to the general (Plano Clark & Creswell, 2008). The researcher makes observations and finds underlying themes in the data and then these findings are subjected to more testing for further clarifications of the theme (Blaikie, 2007). The inductive process is as follows: observations, findings and then theory (Bryman & Bell, 2007).

In this study the researcher follows an objective assumption, a positivist approach and a deductive logic because using a scale to gather information from a large number of people to be analysed and then generalised to some degree requires a positivist approach. The positivist approach is used by many researchers in service quality and is predominately used because of the complex nature of service quality (Schembri & Sandberg, 2002). Parasuraman, Zeithaml & Berry (1998) who originally developed the SERVQUAL scale, used a positivist approach to measure service quality. Additionally Dabholkar, Thorpe and Rentz, (1996) developed the RSQS scale
using a positivist approach. Ladhari (2008) researched 30 industry specific scales and reported that all of the studies used a positivist approach.

4.4 The Research Methodology and Methods
The knowledge claims, the strategies and the method all contribute to a research approach that tends to be more quantitative, qualitative or mixed (Creswell 2003).

4.4.1. Quantitative Approach
Quantitative Approach is one in which the investigator primarily uses post positivist claims for developing knowledge (i.e., cause and effect thinking, reduction to specific variables and hypotheses and questions, use of instrument and observation, the test of theories), employs strategies of inquiry such as experiments and surveys and collects data on predetermined instruments that yield statistical data (Creswell 2003).

Quantitative research is frequently referred to as hypothesis – testing research. Characteristically, studies begin with statements of theory from which research hypotheses are derived. Then an experimental design is established in which the variables in question (the dependent variables) are measured while controlling for the effects of selected independent variables. Subject included in the study are selected at random is desirable to reduce error and to cancel bias. The sample of subjects is drawn to reflect the population (Newman & Benz 1998).

These procedures are deductive in nature, contributing to the scientific knowledge base by theory testing. This is the nature of quantitative methodology. Because true experimental designs require tightly controlled conditions, the richness and depth of meaning for participant may be sacrificed. As a validity concern, this may be a limitation of quantitative design (Newman & Benz 1998).

4.4.2 Qualitative Approach
Qualitative research is multi method in focus, involving an interpretive, naturalistic approach to its subject matter. This means that qualitative researchers study things in their natural settings,
attempting to make sense of, or interpret, phenomena in terms of the meanings people bring to them (Newman & Benz 1998).

Qualitative approach is one in which the inquirer often makes knowledge claims based primarily on constructivist perspectives (i.e. the multiple meaning of individual experiences, meaning socially and historically constructed, with an intent of developing a theory or pattern) or advocacy/participatory perspectives (i.e. political, issue-oriented, collaborative or change oriented) or both. It also uses strategies of inquiry such as narratives, phenomenology, ethnography, grounded theory studies or case studies. The researcher collects open-ended, emerging data with the primary intent of developing themes from the data (Creswell 2003).

One of the most important and equally difficult decisions to make is whether to use quantitative or qualitative research methods or a mixed method approach. The differences between the qualitative and quantitative approaches are based on the judgment of different authors because both approaches may include different methods (Hanson & Grimmer, 2007). Quantitative research is used to test an objective theory (Creswell, 2009) which usually requires the researcher to collect numerical data and analyse it statistically (Punch, 2001). Therefore, quantitative research usually involves instruments that are used for collecting data (Punch, 2001) and quantitative methods are usually concerned with making generalizations about a population under investigation (Zikmund et al., 2010).

3.4.3 Research Strategy

Research Strategy is a general plan of how researcher goes about answering the research questions that has been set by researcher. It contains clear objectives, derived from research questions specify the sources from which researcher intend to collect data and consider the constraints that researcher have such as access to data, time, location and money ethical issues. (Thornhill et. Al., 2003).

In general, there are two types of quantitative research methods (Creswell, 2007):
1. Survey: This entails the collection of data by using a questionnaire to discover the opinions of a population, based on a sample of the population (Creswell, 2009). Moreover, there are many techniques used to collect data which include interviews, telephone calls and observation (Tharenou, Donohue & Cooper, 2007). The term observation used here has a different meaning to the same term used in the qualitative method; it is the recording of numbers e.g. recording the number of people that cross a particular street. A more detailed discussion of survey is found in the survey section.

2. Experiment: A researcher assigns participants to two groups; a control group and a treatment group that is given the experimental condition. Experiments are performed to test the relationship between cause and effect. This involves a comparison of the control group to the treatment group (Tharenou, Donohue & Cooper, 2007). According to Tharenou, Donohue & Cooper (2007), experiments can be either true experiments or quasi-experimental design.

Based on the objectives of this study, a quantitative method has been used for the development of the grocery store service quality scale. The most widely known service quality model, which is the SERVQUAL, used a quantitative method as the main approach for data collection and the qualitative method was used for generating items for the service quality scale (Chen et al., 2007). According to Ladhari (2008) qualitative methods make it difficult to find small differences that quantitative research might be able to find. Additionally, qualitative methods usually use small samples which make the findings not generalisable to the population (McDaniel & Gates, 2006).

As discussed earlier surveys, using questionnaires are the most common way to collect data in service quality studies. This is explained in detail below.

4.4.3.1 Survey

According to Zikmund and Babin (2010, p.64) a survey is a research technique in which responses is collected through structured instrument from a sample in some form or the behaviour of respondents is observed and described in some way. Generally surveys are linked to deductive logic and are a regular method of collecting data in management research by employing a questionnaire that collects data from a sample then statistically analyses the data (Saunders,
Lewis & Thornhill, 2007). Surveys are usually used to collect quantifiable data from respondents to measure, examine, analyse and generalize the findings. Furthermore, it has become accepted as a scientific and accurate way of collecting data to quantify gathered information, even though some aspects of the survey might be qualitative (Zikmund, 2003). The qualitative method usually uses small samples which mean the findings cannot be generalized to the population (McDaniel & Gates, 2006).

Surveys usually record the perception of respondents about opinions, attitudes or beliefs and it is a way of understanding consumer preferences (Black, 1993; Fowler, 2002). Surveys are an inexpensive and efficient way of gathering information about the population or a sample of it (Hague, 2002). This information could be collected from respondents by making telephone calls, sending the questionnaire by post, face to face or using other methods of data collection (Hague, 2002). As the main reason for having a questionnaire is to obtain information from the respondent, it should be logical and follow a certain order that allows respondents to record facts, comments or attitudes (Hague, 2002).

A popular method of collecting data for surveys is using Likert scales, these are used for measuring attitudes which require respondents to choose a statement from a number of statements that range from 'strongly agree' to 'strongly disagree' (Zikmund, 2003; Saunders, Lewis & Thornhill, 2007). The respondent usually chooses a response from a set of five statements where each response is assigned a weight which allows the researcher to perform statistical analysis (Zikmund, 2003). The choice that the respondent makes means that they agree with the statement they have chosen which allows them to express their feelings (Zikmund, 2003). Moreover, the questionnaire should be organized in a way that places similar questions in the same category to make it easier for the respondent to follow (Saunders, Lewis & Thornhill, 2007).

The researcher used a Likert scale that adopted a 5 point scale for his study, instead of a 7 point Likert scale because 5 point scales reduce the level of frustration among respondents, and increases the rate and quality of the responses (Buttle, 1996; Prayag, 2007). Likert scales are a dominant method used for measuring service quality; twenty five out of the thirty studies that
developed industry specific scales, used a Likert scale to measure service quality (Ladhari, 2008). Another study that reviewed 19 service quality models found that 10 used quantitative methods which all used Likert scales (Seth, Deshmukh and Vrat, 2005). According to Caro and Garcia (2008) the number of items on the service quality scale depends on the context of the research. A Likert scale was used in this study which asked customers to best describe the way they feel about the service quality found during shopping in grocery stores as: 1 strongly disagree, 2 disagree, 3 neither agree nor disagree, 4 agree and 5 strongly agree.

4.5 Sample
A sample is taken from the target population being researched. If the sample is adequate it will have the same characteristics of the population (Zikmund, 2003) and the findings are usually used to make conclusions about the population (Field, 2009). Moreover, sampling techniques are methods that are used to select a sample from the population by reducing it to a more manageable size (Saunders, Lewis & Thornhill, 2007). According to de Leeuw, Hox and Dillman (2008), these sampling techniques are used when inferences are made about the target population.

There are two main sampling techniques (Zikmund, 2003; Saunders, Lewis & Thornhill, 2007):

1. Probability Sampling
It refers to the possibility that each object in the target population could be selected and usually the chance of selecting one object is equal to the chance of selecting any other object. This method is usually used with surveys and generalizations about the population are obtained from the sample. Probability sampling has four important elements. First, the sampling frame which refers to the population under investigation. Second, to decide on an appropriate sample size, the larger the sample the more accurate are the generalizations about the population. Third, high response rate is important because the higher the response rate the more representative the sample is of the population. Fourth, select the best sampling technique for the study from the following:
Simple Random: Each unit in the population has the same chance of being selected. This could be done by assigning numbers to each object in the population and then drawing a number at random.

Systematic Sampling: The starting point is randomly selected and then every 3rd, 4th item etc. is selected.

Stratified Sampling: The population is divided into strata (groups) that have similar characteristics and then samples are drawn from each group.

Cluster Sampling: The population is divided into groups (clusters) which are similar to the stratified sampling, but in this method there are —natural groupings of the population.

Multi Stage Sampling: This is also called multi stage cluster samples which are similar to the cluster sampling but in this instance the samples are taken at different stages from the different clusters.

2. Non-probability Sampling

It refers to objects that are selected based on the judgment of the researcher and usually generalizations from this method are possible, although not by using statistical techniques. According to de Leeuw, Hox and Dillman (2008) statistical inferences are not appropriate when non-probability sampling is the technique used. The non-probability sampling also takes many forms (Zikmund, 2003):

Convenience sampling (haphazard sampling): This is the selection of the units that are most convenient for the researcher but makes it difficult to discern if the chosen sample is representative of the population.
Quota sampling: The population is divided into different groups that have similar characteristics; the units are then taken from each subgroup. This is a non-random technique that is usually used for interviews.

4.5.1. Sampling for this Study: In this study Non-probability Quota sampling method was used for the development of GSSQS scale.

4.5.2 Sample size
Sample size has an effect on how the sample findings accurately represent the population (Bums & Bush, 2010). The larger the sample is, the more likely that the generalisations are an accurate reflection of the population (Saunders, Lewis & Thornhill, 2009). Sample sizes depend on factors such as the time and money available to collect the data (Hair 2006); they also depend on the statistical analysis used in the study (Saunders, Lewis & Thornhill, 2009). According to Hair (2006), small or very large samples have a negative impact on the statistical tests because either the sample is either not big enough to make generalizations or too big to reach any conclusions. Tabachnick and Fidell (2007) established that a sample size of 300 is adequate for factor analysis and for regression analysis a sample size of $N \geq 50 + 8* M$ is adequate where $M$ is the number of independent variables. In general, there has been an understanding among authors of statistical books that the larger the sample the more appropriate it is for factor analysis (Pallant, 2007). Hair (2006) suggested that a sample size larger than 100 is needed for factor analysis and as a general rule of thumb, the observations should be 5 times the number of variables. The number of variables in this study before factor analysis is 58 questions which suggests that the sample size should be larger than 500. There are 6 independent variables in this study; a sample size of $50 + 8*6 = 98$ is adequate for regression analysis.
4.5.2.1. Sample size for this study:

In the first phase of study, total 300 hundred questionnaires were distributed out of 218 responses were found usable.

Table 4.3: Sample Plan- phase 1: Geographic Area and Type of Store wise

<table>
<thead>
<tr>
<th>Location</th>
<th>Departmental Store</th>
<th>Discount Store</th>
<th>Super Market/ Hyper Market</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delhi</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Ghaziabad</td>
<td>20</td>
<td>20</td>
<td>35</td>
<td>75</td>
</tr>
<tr>
<td>Noida</td>
<td>20</td>
<td>20</td>
<td>35</td>
<td>75</td>
</tr>
<tr>
<td>Gr Noida</td>
<td>10</td>
<td>15</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>Total No.</td>
<td>75</td>
<td>80</td>
<td>145</td>
<td>300</td>
</tr>
</tbody>
</table>

Total no. of usable responses found for analysis- 218

Table 4.4: Completed and usable questionnaires-Phase1: Geographic Area and Type of Store wise

<table>
<thead>
<tr>
<th>Location</th>
<th>Departmental Store</th>
<th>Discount Store</th>
<th>Super Market/ Hyper Market</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delhi</td>
<td>17</td>
<td>18</td>
<td>41</td>
<td>76</td>
</tr>
<tr>
<td>Ghaziabad</td>
<td>14</td>
<td>15</td>
<td>24</td>
<td>53</td>
</tr>
<tr>
<td>Noida</td>
<td>13</td>
<td>15</td>
<td>26</td>
<td>54</td>
</tr>
<tr>
<td>Gr Noida</td>
<td>7</td>
<td>11</td>
<td>17</td>
<td>35</td>
</tr>
<tr>
<td>Total</td>
<td>51</td>
<td>59</td>
<td>108</td>
<td>218</td>
</tr>
</tbody>
</table>
In the second phase of study total 200 hundred questionnaires were distributed out of 177 responses were found usable.

**Table 4.5:** Sample Plan- Phase 2: Geographic Area and Type of Store wise

<table>
<thead>
<tr>
<th>Location</th>
<th>Hyper Market</th>
<th>Super Market</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delhi</td>
<td>35</td>
<td>40</td>
<td>75</td>
</tr>
<tr>
<td>Ghaziabad</td>
<td>25</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>Noida</td>
<td>25</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>Gr Noida</td>
<td>12</td>
<td>13</td>
<td>25</td>
</tr>
<tr>
<td><strong>Total No.</strong></td>
<td><strong>97</strong></td>
<td><strong>103</strong></td>
<td><strong>200</strong></td>
</tr>
</tbody>
</table>

Total no. of usable responses found for analysis- 177

**Table 4.6:** Completed and usable questionnaires- Phase 2: Geographic Area and Type of Store wise

<table>
<thead>
<tr>
<th>Location</th>
<th>Hyper Market</th>
<th>Super Market</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delhi</td>
<td>32</td>
<td>34</td>
<td>66</td>
</tr>
<tr>
<td>Ghaziabad</td>
<td>21</td>
<td>23</td>
<td>44</td>
</tr>
<tr>
<td>Noida</td>
<td>23</td>
<td>23</td>
<td>46</td>
</tr>
<tr>
<td>Gr Noida</td>
<td>10</td>
<td>11</td>
<td>21</td>
</tr>
<tr>
<td><strong>Total No.</strong></td>
<td><strong>86</strong></td>
<td><strong>91</strong></td>
<td><strong>177</strong></td>
</tr>
</tbody>
</table>
For the main study a total of 900 questionnaires were distributed out of which 703 usable responses were used for the main study which exceeded the minimal requirement for factor analysis and regression analysis.

Total no. of Questionnaires Distributed - 900

**Table 4.7: Sample Plan- Phase3: Geographic Area and Type of Store wise**

<table>
<thead>
<tr>
<th>Location</th>
<th>Hyper Market</th>
<th>Super Market</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delhi</td>
<td>125</td>
<td>125</td>
<td>250</td>
</tr>
<tr>
<td>Ghaziabad</td>
<td>75</td>
<td>75</td>
<td>150</td>
</tr>
<tr>
<td>Noida</td>
<td>80</td>
<td>70</td>
<td>150</td>
</tr>
<tr>
<td>Gr Noida</td>
<td>50</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Faridabad</td>
<td>50</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Ballabgarh</td>
<td>10</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Gurgaon</td>
<td>65</td>
<td>65</td>
<td>130</td>
</tr>
<tr>
<td>Total No.</td>
<td>455</td>
<td>445</td>
<td>900</td>
</tr>
</tbody>
</table>
Total no. of usable responses found for analysis- 703

**Table 4.8:** Completed and usable questionnaires- Phase 3: Geographic Area and Type of Store wise

<table>
<thead>
<tr>
<th>Location</th>
<th>Hyper Market</th>
<th>Super Market</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delhi</td>
<td>97</td>
<td>101</td>
<td>198</td>
</tr>
<tr>
<td>Ghaziabad</td>
<td>55</td>
<td>41</td>
<td>96</td>
</tr>
<tr>
<td>Noida</td>
<td>66</td>
<td>59</td>
<td>125</td>
</tr>
<tr>
<td>Gr Noida</td>
<td>31</td>
<td>44</td>
<td>75</td>
</tr>
<tr>
<td>Faridabad</td>
<td>40</td>
<td>42</td>
<td>82</td>
</tr>
<tr>
<td>Ballabgarh</td>
<td>5</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>Gurgaon</td>
<td>60</td>
<td>54</td>
<td>114</td>
</tr>
<tr>
<td><strong>Total No.</strong></td>
<td><strong>354</strong></td>
<td><strong>349</strong></td>
<td><strong>703</strong></td>
</tr>
</tbody>
</table>

**4.6 The Research Objectives**

The first objective of this study is to test the applicability of existing scale- RSQS on grocery store’s service quality. The second objective of the research is to identify/ develop new dimensions and attributes develop the extended instrument. Third objective of this study is to develop a scale to measure Service Quality for Grocery Retail Stores. Fourth objective of this study is to measure any differences between the overall service quality by demographic dimensions. Based on the literature review and to meet the objectives of this study the main research question is:

How to evaluate retail service quality in a grocery retail store and its influence on customers?
The following supporting questions emerged to answer the overall research question.

1. **How should service quality perception be measured in a grocery retail store using a context-specific scale?**

2. **Is there an association between service quality dimensions and overall service quality?**

3. **Is there a difference between service quality dimensions by demographic information?**

The relationship between service quality dimensions and overall service quality and demographic information is in line with previous research as evident in the literature review chapter. The first research question will be answered by testing the reliability and validity of the scale to determine whether it is suitable for measuring service quality in a grocery retail store. The second and third research questions are answered by the following hypotheses:

### 4.6.1 The Research Hypotheses

The following hypotheses will answer the second and third research questions and the objectives of the research study.

In order to know the association between dimensions of GSSQS and overall service quality, the first set of hypotheses has been formulated.

**First set of hypotheses**

*H1* - The modified service quality dimensions of GSSQS have a positive association with overall service quality.

*H1.1.* There is no association of the dimension 1 (Grocery Service) with overall service quality.

*H1.2.* There is no association of the dimension 2 (Discounts & Promotions) with overall service quality.

*H1.3.* There is no association of the dimension 3 (Courteousness) with overall service quality.
H1.4. There is no association of the dimension 4 (Problem Solving) with overall service quality.

H1.5. There is no association of the dimension 5 (Policy) with overall service quality.

H1.6. There is no association of the dimension 6 (Physical Evidence) with overall service quality.

In order to know the significance difference in GSSQS scale dimensions by demographic information, the second set of hypotheses has been formulated.

Second set of hypotheses

H2- There is a significant difference in GSSQS scale dimensions by demographic information.

H2.1. There is no significant difference in GSSQS scale dimensions by the customer’s age groups.

H2.2. There is no significant difference in GSSQS scale dimensions by customer’s gender.

H2.3. There is no significant difference in GSSQS scale dimensions by customer’s marital status.

H2.4. There is no significant difference in GSSQS scale dimensions by customer’s Occupation.

H2.5. There is no significant difference in GSSQS scale dimensions by market type of markets visited (Hyper-market vs. Super market).

4.7 Method of Data Analysis

Data analysis comes after the data has been collected (Field, 2009) to make sense of the study and reach certain findings. This section will present the different techniques used for data analysis by the researcher in this study.

4.7.1 Factor Analysis

One of the major importances of factor analysis is to summarize the data to be more manageable without losing any of the important information therefore making it easier to test theories (Field, 2009; Johnson and Wichern, 1998; Tabachnick and Fidell, 2007). There are three main reasons for using factor analysis (Field, 2009): to develop a scale to measure service quality of Grocery

66
retail stores, to reduce the variables to a manageable size and to have a better understanding of the variables.

According to Cooper and Schindler (2008) factor analysis is a technique used for specific computational techniques. These factors, also called latent variables, aim to measure things that are usually hard to measure directly, such as attitudes and feelings (Field, 2009). This is a way to explain the relationships among variables by combining them into smaller factors (Coakes and Steed, 2001; Zikmund, 2003). The scales usually start with many questions, and then by using factor analysis are reduced to a smaller number (Pallant, 2007). These reduced results are then used for other analysis such as multiple regression analysis (Pallant, 2007).

There are two methods of factor analysis observed by Kinnear and Gray (2010), the exploratory factor analysis and the confirmatory factor analysis. The purpose of the exploratory factor analysis is to find the number of factors that explain the correlations; while in the confirmatory factor analysis the researcher predicts the number of factors with specific loading (Kinnear and Gray 2010). Another important point to consider when performing factor analysis is factor loadings. Factor loadings are the correlation of the variable with the factor. When the loading is clear then the interpretations of the factors become easier (Zikmund et al., 2010). Some variables have a loading or correlation with more than one factor. The mathematical technique for simplifying the results of the factor analysis results is called factor rotation (Zikmund et al., 2010).

The most common method of factor analysis is the principal component (Cooper & Schindler, 2008; Kinnear & Gray, 2010) and the most common method of factor rotation is the varimax rotation (Kinnear & Gray, 2010; Zikmund et al., 2010). Principal component technique looks at the correlation of different variables to reveal the relationship between them, and then reduces the variables by empirically summarizing them or combining them into a small number of factors under common themes (Tabachnick and Fidell, 2007). Factor rotation is used as a method to interpret the factors by showing the variables that group together (Pallant, 2007). There are two tests performed to ensure that the data is suitable for factor analysis, the Kaiser-Meyer-Olkin
(KMO) measure for sampling adequacy and the Bartlett’s test of sphericity (Pallant, 2007). The KMO value is low if it is between 0.5 and 0.7 and excellent if it is above 0.90 (Field, 2009). Factors with an Eigen value of 1 or greater are usually retained (Field, 2009).

The researcher will use principal factor analysis using principal component analysis with varimax rotation to reduce and to analyze the data collected from the modified instrument. Factor analysis is heavily used for service quality questionnaires and, according to Gilbert and Veloutsou (2004), this technique has been adopted by almost one sixth of the authors of journal articles over the past thirty years. Moreover, the method used by Gilbert et al. (2004) and Akbaba (2006) in their studies is the principal components and varimax procedure.

4.7.2 Regression Analysis
Regression analysis is used to find the relationship between one dependent variable and one or more independent variables and has become popular in many research areas (Hair et al., 1998; Tabachnick & Fidell, 2007). Regression is another way to determine the association between variables; this is similar to bivariate correlation as it assumes that there is a link between the dependent and independent variable (Zikmund, 2003). Regression analysis is called simple regression analysis when there is only one independent variable and is called multiple regression analysis when there is more than one independent variable (Robson, Pemberton & McGrane, 2008). Multiple regression analysis is more complex than correlation and is used to find the ability of a set of independent variables in predicting the dependent variable (Pallant, 2007).

According to Zikmund, Babain & Griffin (2010) simple regression analysis and correlations are mathematically the same in many respects; however the correlation is an interdependence technique and regression is a dependence technique. The interdependence technique does not differentiate between dependent and independent variables; whereas the dependence technique distinguishes between dependent and independent variables.

The R square ranges from 0-1 and this shows how much of the dependent variable is explained by the independent variables (Burns & Bush, 2010). The higher the R square the stronger the
association between the dependent variable and the independent variable (Burns & Bush, 2010). When performing regression analysis normality is not always required when analysing variables. However the results are slightly improved if the data is normally distributed (Tabachnick & Fidell, 2007).

To perform multiple regression analysis, these assumptions must be adhered to:

- **Sample size:** The sample size has to be large enough for the results of the regression analysis to be meaningful (Tabachnick & Fidell, 2007). The sample size must be $N \geq 50 + 8*m$ for testing multiple correlations, where $m$ is the number of independent variables.

- **Multicollinearity:** Multicollinearity should not exist because multicollinearity problems weaken the regression model (Pallant, 2007). Multicollinearity exists when there is a strong relationship between the independent variables. The independent variables must not correlate highly with each other; $r$ value is preferred to be below 0.7 and an $r$ value of above 0.9 indicates strong correlation which weakens the regression model (Pallant, 2007). To perform the regression analysis there should also be some correlation between the independent variables and the dependent variable, it should not fall below 0.3 (Pallant, 2007). To test multicollinearity the tolerance level must be greater than .10 and the VIF must be less than 10.

- **Outliers:** This occurs when there is a value that is extreme for one variable or different combinations that are not logical (Tabachnick & Fidell, 2007). According to Tabachnick and Fidell (2007) there are four reasons for having outliers: the data is entered incorrectly, missing values are not detected properly, a case that is entered that does not belong to the population and the population has many extreme values. Mahalanobis is used to detect outliers.

Regression analysis is used for service quality studies that involve the attitudes and perceptions of consumers or the decisions that consumers make about products (Hair et al., 1998). A study that measured retail service quality by using the RSQS in discount stores in Korea and the US
used regression analysis to find the relationship between service quality, behavioral intentions and satisfaction (Kim & Jin, 2001). The dimensions of the service quality were tested as the independent variable while the behavioral intentions and satisfaction were tested as the dependent variables. Another study that used regression analysis measured service quality in Islamic banks by testing the relationship between the overall service quality as the dependent variable and the service quality dimensions as the independent variable (Al-Tamimi & Al-Amiri, 2003). In a study by (Mehta, Lalwani & Han, 2000) regression analysis was used to find the significance of the RSQS dimensions in explaining the variance in the overall perception of the retail environment. Akbabba (2006) used regression analysis to test the relationship between the overall service quality against the service quality dimensions. In this study multiple regression analysis was used because there is more than one independent variable. Overall service quality was the dependent variable and the service quality dimensions were the independent variables.

4.8 The Reliability and Validity of the Scale

Reliability and validity of the service quality scale is important for the measurement of service quality and for obtaining meaningful results. Furthermore, reliability is more important when the questionnaire is a Likert-type because there are many variables testing the concept. A questionnaire is considered reliable if it gives similar results when repeated (Hair, 2003).

4.8.1 Reliability

Reliability is an indication of how consistent the findings are based on the method of data collection and analysis (Saunders, Lewis & Thomhill, 2007). According to Zikmund and Babin (2010b, p.334) reliability is an indicator of a measure's internal consistency. The most common method for testing the internal consistency of a scale for reliability is the Cronbach alpha coefficient (Hair et al., 1998; Pallant, 2007). The Cronbach alpha coefficient ranges from 0 to 1 with a minimum of 0.6 while other studies suggest that anything above 0.7 suggests high levels of internal reliability (Hair et al., 1998). Hair et al. (1998) stated that reliability is a test of the amount of consistency between the different measurements of an individual's response, with the purpose of making sure that these responses are consistent and similar over a period of time and across situations. Furthermore, if the same test was conducted under the same circumstances,
then the researcher should find similar results which is called the test-retest method (Zikmund & Babin, 2010). However, this method has two problems; the first has to do with the test being longitudinal, this indicates that the results for the second test might be influenced by the results of the first test. The second problem is regarding the homogeneity of the test, meaning that the questions might need to be altered slightly (Zikmund, 2003). To check internal consistency, which means that the measure is homogenous, there are several methods such as the split half method and the equivalent form method. In the split half method, results are split in half and compared to see if the results are similar. With the equivalent form method, the researcher is required to develop two similar scales to test the same thing and then compare the scales to discover if there are high correlations between the two scales. Therefore, if the scales have a high correlation, the test is reliable (Zikmund, 2003). For the purpose of this research the researcher will use Cronbach alpha coefficient, the most common method for testing reliability, and 0.7 will be used as the minimal accepted level. Many studies have used reliability to test their modified service quality scale that ranged from 0.6 to 0.96 (Dabholkar, Thorpe & Rentz, 1996; Sureshchandar, Rajendran & Anantharaman, 2002; Jabnoun & Khalifa, 2005; Akbaba, 2006; Caro & Garcia, 2007; Chowdhary & Prakash, 2007).

4.8.2 Validity

According to Zikmund and Babin (2010b, p.335) validity is —the accuracy of a measure or the extent to which a score truthfully represents a concept. Validity is concerned with the test being capable of testing what it was designed for, which is not as simple as it seems (Hair, 2003). Normally researchers refer to three types of validity for the testing of a scale (Saunders, Lewis & Thornhill, 2007; Zikmund & Babin, 2010b). Zikmund and Babin (2010b) describe the three types as follows:

- **Face validity or content validity**: It refers to an agreement between experts that the scale measures what it is intended to and seems to be a good reflection of the scale (Zikmund & Babin, 2010b). To verify face validity, the researcher reviews the literature to find similar scales used by experts in the field and where possible, consults with the experts (Parasuraman, Zeithaml & Malhotra, 2005).
Criterion validity: It refers to checking the scale to ascertain whether it can give accurate predictions, usually of the dependent variable. Moreover, it could refer to obtaining similar results using different measurements. Criterion validity could be classified as either predictive validity or concurrent validity. The classification depends on the time in which the scale and criterion validity are correlated. Concurrent validity is when the new measure and the validity are carried out at the same time and the measure turns out to be valid. Predictive validity is when the measure predicts a future result. A number of studies that have measured predictive validity include Sureshchandar, Rajendran and Anantharaman (2002); Jabnoun and Khalifa (2005), and Parasuraman, Zeithaml & Malhotra (2005).

Construct validity: It refers to the analysis of the data supporting a part of the study’s hypothesis or the scale answering some of the research questions. Construct validity includes convergent validity and discriminant validity. Convergent validity: to have convergent validity, the scores for the category 'excellent' must be higher than the category 'very good' for each of the dimensions (Parasuraman, Zeithaml & Berry, 1988; Aldlaigan & Buttle, 2002). The scores for the category 'good' must also be higher than the scores for the category 'poor' and so on. Discriminant validity shows that the measure is unique in some way. Studies that have measured construct validity include Parasuraman, Zeithaml & Malhotra (2005); and Caro & Garcia (2007). Using a scale outside of its cultural context runs the risk of the instrument not measuring what is was intended to measure. Therefore, it is crucial to test the instrument for construct validity before using it in a different country (Cui, Lewis & Park, 2003).

4.9 Scale Modification
Evidence from past research suggests that using a generic scale to measure service quality across industries is not suitable without modification (Akaba, 2006; Caro & Garcia, 2008; Ladhari, 2008); therefore, the more specific the measure is, the more valuable the potential information could be (Karatepe, Yavas & Babakus, 2005; Ladhari, 2008). It is common for scales to be
modified when measuring service quality, and this has been done by many researchers who felt the need for industry specific measures (Karatepe, Yavas & Babakus, 2005; Chowdhary & Prakash, 2007) and culture specific measures (Cui, Lewis & Park, 2003; Karatepe, Yavas & Babakus, 2005; Prayag, 2007). To customize a scale or adapt questions that have already been used in other studies is more efficient than using questions that have not been tested, given that they are adequate for collecting the data needed by the researcher (Saunders, Lewis & Thornhill, 2007).

The SERVQUAL scale extends across many industries and cultures. For this reason it becomes a weaker instrument that is used as a base for developing new service quality scales in different cultures (Gaur & Agrawal, 2006). The SERVQUAL and the RSQS are two scales that are most frequently used to measure retail service quality. However, they are not suitable for use in other cultures without modifications (Gaur & Agrawal, 2006). A study conducted in Singapore set out to discover what the most appropriate scale was, RSQS or SERPERF, which only measures the perception level of service, in supermarkets and electronic retailers. They found that RSQS is superior in an environment where there are more goods and less services e.g. supermarkets, while SERVPERF is more appropriate for environments where services are more important than goods e.g. electronic retailers (Mehta, Lalwani & Han, 2000). Based on the fact that scales need to be industry specific and culturally sensitive, Grocery stores have more goods than services, the researcher decided to modify a scale based on the RSQS. The scale modification steps are discussed further.
Figure 4.1: Scale modification steps

Source: (Churchill, 1979; Parasuraman, 1989; Ahmad et al., 2009)
Figure 4.1 above shows the scale modification steps proposed by Churchill (1979), Parsuraman (1989) and later proposed by Ahmad et al (2009). The steps taken in this research are summarized below:

- **Specify domain of construct:** Based on an extensive review of the literature it was concluded that the most referred scale for measuring retail service quality is the RSQS. The first phase of the study was to check the reliability and validity of the scale in India especially for ‘Grocery Retail Stores’. RSQS, which has been already studied by many other authors like Subhashini Kaul on garment stores, proved that it is not valid in India. Thus we tried to test it for grocery retail stores, in order to develop a new/modified scale for the same.

- **1st Phase:** Testing of RSQS on grocery Retail Stores: Data was collected using RSQS instruments. Non probability Quota sampling was used for data collection from Delhi NCR region. The data was run to test the reliability and validity of the scale, which was found negative.

- **II**nd Phase: Identification of new variables to develop extended instrument for the purpose of developing new scale for grocery retail stores (Grocery Store Service Quality Scale –GSSQS): An exploratory customers’ Top of the Mind (TOM) survey was conducted to find out the new attributes and dimensions. In addition to TOM survey, the experts from academics and corporate were consulted to know the various important factors leading the customers’ purchasing behavior in grocery retail stores. Some other attributes were included which were referred in research papers of previous researchers in context of grocery retailing. Based on these three surveys, the extended instrument was prepared with overall 58 items under 10 dimensions.

- **III**rd Phase: Development of a New/modified scale GSSQS: Non Probability Quota sampling was used in this study which is the preferred method for surveys (Saunders, Lewis & Thornhill, 2009) and the type of Non probability sampling used was Quota sampling as suggested by Dabholkar, Thrope & Rentz (1996) for their study. To check the relationship between dependent and independent variables, regression analysis was done.
- **Purify measure**: to purify the scale exploratory factor analysis was performed for reduction and grouping of items of the scale. A principal component analysis with varimax rotation was used and this reduced the 58 item instrument to 25 items scale.

- **Assess reliability**: the reliability of the scale was tested using coefficient alpha which showed the scale highly reliable.

- **Assess validity**: the validity of the scale was tested using three types of validity, face validity, criterion validity and construct validity.

### 4.10 Conclusion

This chapter presents an overview of the methodology used in this study which was selected to support the objectives of the study. The whole research was completed in three phases.

In the first phase, to test the reliability and validity of the RSQS scale, a quantitative survey was conducted. That study did not prove the reliability and validity of RSQS when measuring service quality of a grocery retail store.

In the second phase, an extensive TOM survey was done for the identification of new variables and attributes, to know what the prime factors in the mind of customers are when they are choosing a grocery retail store for purchasing grocery for their household. In addition to TOM survey, the experts from academics and corporate were consulted to know the various important factors leading the customers' purchasing behavior in grocery retail stores. Some other attributes were included which were referred in research papers of previous researchers in context of grocery retailing. Based on these three surveys, the extended instrument was prepared with overall 58 items under 10 dimensions.

The extended instrument was used in the last phase of study. To purify the scale exploratory factor analysis was performed for reduction and grouping of items of the scale. The reliability of the scale was tested using coefficient alpha which showed the scale highly reliable. The validity of the scale was tested using three types of validity, face validity, criterion validity and construct validity.