CHAPTER - 3

RESEARCH DESIGN AND METHODOLOGY

The validity of any research depends on the systematic method of collecting the data and analyzing the same in a logical and sequential order. The study is intended to cover the factors that influence the investors in mutual fund selection. Data for this investigation were collected from primary as well as secondary sources.

The chapter presents the methodology that was adopted by the researcher in conducting the study titled “DETERMINANTS OF MUTUAL FUND SELECTION BY RETAIL INVESTORS”. The methodology includes the following aspects:

3.1 Period of study
3.2 Selection of area for the study
3.3 Method of data collection
   (i) Primary data collection
   (ii) Secondary data collection
3.4 Testing validity and reliability of the instruments
   (i) Validity
   (ii) Reliability
3.5 Sample size selection and sampling techniques
3.6 Statistical tools used for analysis
3.7 Concepts used in the study

3.1 Period of study- The study was conducted during the period June 2011 - May 2013

3.2 Selection of area for the study

Coimbatore District is one of the fast-growing districts of the state of Tamil Nadu in India. The NCAER HOUSEHOLD SURVEY REPORT JULY 2011 reveals that more than 90 per cent of the investors are individual investors and 40.8 per cent of the households invest in mutual funds. Coimbatore stands 17th place among 54 cities with 1.17 per cent of total investors. This made the
researcher to focus on the Coimbatore district. During the study period (November 2012 to March 2013) the district had only two revenue divisions, namely Coimbatore and Pollachi with eight taluks, namely Coimbatore (North), Coimbatore (South), Mettupalayam, Annur, Sulur, Kinathukadavu, Pollachi and Valparai.

3.3 Method of data collection

(i) Primary Data

The data needed for the study is primary in nature. In order to fulfill the objectives, the study was undertaken by using a well-framed questionnaire that was duly filled in by the respondents.

(a) Research instrument

The study was based on the primary data collected from the respondents with the help of a structured questionnaire. The questionnaire is comprised of five sections. The first section deals with the profile of the investors, their savings and investment information. The second section focuses on financial literacy and awareness level of investors about mutual funds. The responses were measured with 17 statements using 3-point scaling technique (agree, disagree and no idea). The third section has 17 statements to understand the personality traits of the investors and their risk profile. This is based on “Big five” factors in personality traits developed by Botwin and Buss, 1989; Conley, 1985; Eysenck, 1986; Cattle, 1994; Hogan, 1986; MMPI; Myres & Mc Cauley, 1985; Norman, 1967; Goldberg, 1996; Saucier, 1997;) The responses of the investors were measured with five-point Likert's scaling technique (strongly disagree, disagree, neutral, agree and strongly agree). The fourth section lists the different qualities of the mutual fund; the responses of investors were measured with 35 statements using five-point Likert's scaling technique (essentially important, very important, somewhat important, somewhat unimportant and not important). The last portion deals with the source of information about mutual fund, investment method, the preferred mutual fund schemes, satisfaction level of mutual fund investors and problems faced by them.
A pilot questionnaire was first administered to small samples of 50 mutual fund investors during 1st July 2012- 30th September 2012. After pre-testing, necessary modifications were made in the questionnaire to fit in the track of the present study.

The final collection of data was initiated in the month of November 2012 and completed by the end of May 2013 with a response rate of 90 per cent.

(ii) Secondary Data

The primary data were supplemented by secondary sources of data. The secondary data pertaining to the study were gathered from the records published by mutual fund companies. Further, the secondary data were also collected from various leading journals, reports, books and websites to obtain the pertinent literature on financial literacy, mutual fund selection and personality traits.

3.4 Testing validity and reliability of the instrument

(i) Validity

In order to ascertain the validity of the instrument, the questionnaire so drafted was circulated among some research experts, mutual fund investors, agents and research scholars for a critical review with regard to wording, format, sequence and the like. The questionnaire was re-drafted in the light of their comments.

(ii) Reliability

Reliability refers to the level of dependability of the items in the research instrument. The reliability was measured by using the Cronbach- Alpha Coefficient. The correlation coefficient attained from the questionnaire on investment objective was 0.721 and current attitude towards investment type was 0.715. The correlation co-efficient attained for financial literacy and mutual fund conceptual awareness level which includes 17 items was 0.702, personality traits assessment with 17 items was 0.727 and mutual fund qualities which includes 35 items was 0.955. The result indicated that the instruments were reliable as the reliability coefficients were more than 0.700.
3.5 Sample size selection and sampling techniques

The universe of the study consists of all retail investors in each of the eight taluks of the Coimbatore district. From each of the eight taluks, 100 respondents were selected using non-probabilistic convenient and judgment sampling method and the total sample size was 800 retail investors initially. A well-structured questionnaire was issued to all the 800 retail investors, Out of which 80 questionnaires were found to be incomplete and were rejected. Finally 720 questionnaires which were found to be valid in every respect were considered for the study. Thus the response rate was at 90 per cent. The sample for the study is 720 individual investors for primary data collection.

3.6 Statistical tools used for analysis

The analysis was carried out using IBM Statistical Package for the Social sciences (SPSS) 21 version for windows. The difference in the extent of selecting the mutual funds between the different types of respondents based on their age, gender, educational qualification, marital status, family size, employment status, source of income annual income, number of earning members in the family, annual savings and investment time horizon was studied by means of percentages, averages, ranges, standard deviation, two-way classification tables, Chi-Square test, correlation analysis and multiple regression analysis. Further Discriminant analysis, cluster analysis, kruskal-wallis test, factor analysis and Structural Equation Model (SEM) have been used appropriately. In addition, Henry Garrett Ranking Technique has been used to rank the problems faced by the investor while investing in mutual fund.

(i) Chi-Square test

The chi-square technique was applied to find a significant association between independent variables and investment objectives and demographic variables and current attitude towards investment types.
The formula used for the Chi-square ($\chi^2$) test is:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Degree of Freedom (D.F.) = (c-1)(r-1)

Where, $O =$ Observed frequency,
$E =$ Expected frequency,
$C =$ Number of Columns,
$r =$ Number of Rows.

(ii) **ANOVA (Analysis of Variance)**

To study the significant difference between demographic variables of individual investor, their level of financial literacy and level of influence of mutual fund qualities towards mutual funds selection ANOVA technique was applied. It is also applied to understand the significant difference between various clusters of mutual fund investors and choice of investment avenue. Formula used for ANOVA is given below:

$$SS_{Total} = \sum_{j=1}^{p} \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2$$

$$SS_{between} = \sum_{j=1}^{p} n_j (\bar{x}_j - \bar{x})^2$$

$$SS_{within} = \sum_{j=1}^{p} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

$$F = \frac{S1^2}{S2^2}$$

Where $S1^2 =$ between column variance
$S2^2 =$ within column variance

(iii) **Post Hoc test**

Post Hoc tests are designed for situations in which the researcher has already obtained a significant F-test with a factor that consists of three or more
means and additional exploration of the differences among means is needed to provide specific information on which means are significantly different from each other. To explore all possible pair-wise comparisons of means comprising a factor, the post hoc test has been applied. The Least Significant Differences (LSD) test between two means is calculated by

$$LSD : t \sqrt{\frac{2MSE}{n}}$$

Where ‘t’ is the critical, tabled value of the t-distribution with the degree of freedom associated with MSE and ‘n’ is the total number of samples.

(iv) Correlation analysis

The correlation analysis is used to find the relationship between the demographic variables and awareness level and level of influence of mutual fund qualities. The coefficient value can range between -1.00 and 1.00. If the coefficient value is +1 it indicates that there exists a positive correlation and the coefficient correlation value is -1 it indicates that there exists a negative relationship. When the coefficient value is found to be 0 means it indicates that there is no linear relationship between two variables.

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

Where:

- $N = \text{number of pairs of scores}$
- $\sum xy = \text{sum of the scores of the paired scores}$
- $\sum x = \text{sum of x scores}$
- $\sum y = \text{sum of y scores}$
- $\sum x^2 = \text{sum of squared x scores}$
- $\sum y^2 = \text{sum of squared y scores}$

(v) Multiple Regression Analysis

It is applied to test the degree of relationship between demographic variables and awareness level of investors, level of influence of mutual fund qualities towards mutual fund selection.
This analysis is adopted where there is one dependent variable that is presumed to be a function of two or more independent variables. In multiple regression, a linear composite of explanatory variables is formed, in such a way that it has maximum correlation with an active criterion variable. The main objective of using this technique is to predict the variability of the dependent variable, based on its co-variance with all the independent variables. It is useful to predict the level of dependent phenomenon through Multiple Regression Analysis models, if the levels of independent variables are given. The formula used for multiple regression is given below:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \ldots + b_kX_k$$

Where, $Y$ is the value of the dependent variables what is being predicted. $X$ is the value of independent variables that is explaining the variables and $Y$ is the value of the Dependent variable ($Y$), what is being predicted or explained

Where $a$ (Alpha) is the Constant or intercept

$b_1$ is the Slope (Beta coefficient) for $X_1$

$S.E . b$ standard error of coefficient $b$

$R^2$ the proportion of the variance in the values of the dependent variable ($Y$) explained by all the independent variables ($X_n$) in the equation together; sometimes this is reported as adjusted $R^2$, when a correction has been made to reflect the number of variables in the equation.

(vi) **Discriminant analysis**

Discriminant analysis is performed to identify the variables discriminating the various method of investment.

(vii) **Cluster analysis**

Cluster analysis is a multivariate statistical technique which groups unknown number of persons / objects / occasions into groups such that the members of each group are having similar characteristics / attributes. The primary objective of Cluster analysis is to define the structure of the data and identify the most similar observations to place them into groups. The different groups to be determined in cluster analysis are not predefined as in Discriminant analysis. The method of
clustering may be either hierarchical or non-hierarchical or both. The outcome of this analysis is much superior when the results from the hierarchical order are used for the analysis along with the non-hierarchical. Thus hierarchical and non-hierarchical techniques should be viewed as complementary clustering techniques rather than as competing techniques. In this study, the researcher has used both hierarchical and non-hierarchical clustering techniques.

Cluster analysis is typically applied to data recorded on interval scale or continuous scaled variables. This analysis is applied to a large set of data which may consist of many variables. Cluster analysis determines internal homogeneity that is similarities that exist among the investors and external heterogeneity that is the differences that exist across different groups of investors personality traits. This analysis helps in grouping the objects or persons based on the variables considered in an analysis or research. This analysis is applied to determine internal homogeneity, and external heterogeneity of investor’s personality traits.

(viii) Kruskal-Wallis test

This test is used to test the null hypothesis that ‘ k’ independent random samples come from identical universe against the alternative hypothesis that the means of these universe are not equal.

It is used to test whether the choice of investment avenues varies with the investment personality.

\[
H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)
\]

Where, \( H \) = Kruskal-Wallis Test statistic
\( N \) = total number of observations in all samples
\( R_i^2 \) = Sum of the ranks assigned

The Kruskal-Wallis test statistic is approximately a chi-square distribution, with k-1 degrees of freedom where \( n_i \) should be greater than 5. If the calculated value of the Kruskal-Wallis test is less than the critical chi-square value, then the null hypothesis cannot be rejected. If the calculated value of Kruskal-Wallis test is greater than the critical chi-square value, then we can reject the null hypothesis and say that the sample comes from a different population.
(ix) **Factor analysis**

Factor Analysis is used to group various items of fund qualities into factors. The purpose of factor analysis is to determine the responses to several statements, which are significantly correlated. Factor Analysis is used to study a complex product or service in order to identify the major characteristics or factors considered important by the respondent. The purpose of factor analysis is to determine the responses to several statements, which are significantly correlated. Factor analysis is applied to assess the significance of the factors that influence the mutual funds selection.

(x) **Structural Equation Modelling (SEM)**

Understanding the way statistical significance is described requires understanding the terminology of the model itself. The structural equation model has graphical display which has boxes and arrows. Boxes denote observed data and the arrows signify assumed causation. In the structural equation model the variable that receives a one-way directional influence from some other variable in the system is termed “endogenous”, or is dependent. A variable that does not receive a directional influence from any other variable in the system is termed as “exogenous” or is independent. When interpreting structural equation model the values attached to one-way arrows (or directional effect) are regression coefficient, whereas two-way arrows (Non directional relationship) are correlation coefficient; Regression coefficients and correlation comprise the “parameters” of the model. The regression coefficient and correlations measures the strength of the relationship between the variable. The regression coefficient of 0.70 or higher indicates a very strong relationship, 0.50–0.69 indicates a substantial relationship, 0.30-0.49 indicates a moderate relationship, 0.10-0.29 indicates a low relationship, 0.01-0.09 indicates a negligible relationship and the value of 0 indicates no relationship.

Besides regression coefficients and correlations, structural equation model also test the overall fit of the model. The narrative analyses use three measures of model fit to determine the overall quality of fit of the model. Another way of describing the model fit is to view this as the test of model significance, thus, when the values of significance are met for the tests all relationships within the model are
significant, and it is their relative strength which decides if there is a relationship or not. Besides testing for model fit, SEMs also provides a measure of multicolinearity. In some cases, the model fits the data well, even though none of the independent variables has a statistically significant impact on the dependent variables. IBM Amos 22.0 has been used to fit the proposed model. It is used to examine a series of dependence relationships simultaneously. It is particularly useful when one dependent variable becomes an independent variable in subsequent dependence relationships. SEM is applied to identify the various dimensions of influence regarding the qualities of the selection of mutual fund.

(xii) Henry Garrett Ranking Technique

The problems faced by the retail investors while investing in mutual fund was analysed using Henry Garrett Ranking Technique. The order of merit given by the respondents was converted into ranks by using the following formula.

$$\frac{100 \times (R_{ij} - 0.5)}{N_j}$$

Percentage Position = The percentage position of each rank thus obtained was converted into scores by referring to the table given by Henry Garrett. Then for each factor the scores of individual respondents were added together and divided by the total number of respondents for whom the scores were added. The factors having highest mean value is considered to be the most important factor. These mean scores for all the factors were arranged in the order of ranks and the inference was drawn.

Where, $R_{ij}$ = Rank given for the $i^{th}$ variable by $j^{th}$ respondents

$N_j$ = Number of variable ranked by $j^{th}$ respondents

With the help of Garrett’s Table, the per cent position can be identified.

3.7 Concepts used in the study

(i) Retail investor

The Ministry of corporate affairs- Investors education and protection fund defines retail investor as “an individual investor who applies or bids for securities/units of or for a value of not more than Rs. 2,00,000.”
(ii) Financial literacy

Financial literacy refers to “providing familiarity with and understanding of financial market and products, especially rewards and risks, in order to make informed choices”.

Financial literacy primarily relates to personal financial education to enable an individual to take effective actions to improve overall well-being and avoid distress in matters that are financial.

(iii) Personality

According to American Psychological Association “Personality” refers to individual differences in characteristic pattern of thinking, feeling and behaving.

(iv) Risk tolerance

“Risk tolerance” is an investing term relating to the amount of market risk, especially the volatility (ups and downs), an investor can tolerate.

(v) Fund qualities

Fund qualities refer to the qualities of a mutual fund, namely fund brand name, transparency factors, expenses ratio, loads, past performance records, brand name of sponsors and agency network.

Mutual fund concepts

(i) Net Asset Value

Net Asset Value is the market value of the assets of the scheme minus its liabilities. The per unit NAV is the net asset value of the scheme divided by the number of units outstanding on the Valuation Date.

(ii) Sale price

It is the price paid at the time of investing in a scheme which is also called Offer Price. It may include a sales load.

(iii) Repurchase price

It is the price at which units under open-ended schemes are repurchased by the Mutual Fund. Such prices are NAV-related.
(iv) Redemption price

The price at which a mutual fund's shares are redeemed (bought back) by the fund. The redemption price is usually equal to the current net asset value per share. It is also called the bid, call or sell price.

(v) Systematic Investment Plans (SIP)

SIP is an investment strategy wherein an investor needs to invest the same amount of money in a particular mutual fund at every stipulated time period.

(vi) Exit load

Mutual funds companies collect an amount from investors when they join or leave a scheme. This fee charged is generally referred to as a 'load'. Exit load is a fee or an amount charged from an investor for exiting or leaving a scheme or the company as an investor.

Big five personality model

The ‘Big Five’ is the commonly used term for the model of personality which describes the five fundamental factors of our personality. The Big Five 'super traits' have been researched and validated by many different psychologists (WT Norman 1963, McCrae & Costa 1987, Brand & Egan 1989, LR Goldman 1990, Botwin and Buss 1989, Conley 1985, Eysenck 1986, Cattle 1994, Hogan 1986, MMPI; Myres & Mc Cauley 1985, Norman 1967, Goldberg, 1996, Saucier 1997, and P Sinclair 1992) and are at the core of many other personality questionnaires.

The Big Five framework of personality traits from Costa & McCrae, 1992 emerged as a robust model for understanding the relationship between personality and various behaviours. The Big Five factors are: Openness (inventive/curious vs. consistent/cautious), Conscientiousness (efficient/organized vs. easy-going/careless), Extraversion (outgoing/energetic vs. solitary/reserved), Agreeableness (friendly/compassionate vs. cold/unkind) and Neuroticism (sensitive/nervous vs. secure/confident). Acronyms commonly used to refer to the five traits collectively are OCEAN, NEOAC, or CANOE.

Keeping the above studies and the study by Chitra et al. (2011) as a base the researcher modified some of the personality trait terms suitable for investment.
Big five personality traits/ five factors model

1. **Openness to experience**

   Describes the breadth, depth, originality, and complexity of an individual’s mental and experiential life. It refers to a person who is imaginative, independent-minded and has divergent thinking. The questionnaire contains statements to know if the investor is curious (Prefer for variety and inventive) or conservative (Prefer for traditional and tend to be conventional).

2. **Extraversion**

   Implies an energetic approach towards the social and material world and includes traits such as sociability, activity, assertiveness, and positive emotionality. The questionnaire contains statements to know if the investor is extrovert (talkative and sociable) or introvert (shy and reserved) or whether he is a risk taker (prefers to take risk) or risk avoider (prefers to avoid risk) and whether the investor is more economic (highly focused on return than risk) or economic investor (thinks return as a part of investment).

3. **Conscientiousness**

   It describes socially prescribed impulse control that facilitates task and goal-directed behaviour, such as thinking before acting, delaying gratification, following norms and rules, and planning, organising, and prioritising tasks. In the current study the questionnaire contains statements to know if the investor is cautious and dependent (planned or organised before any work) or spontaneous and independent investor (the investor is spontaneous).

4. **Agreeableness**

   It contrasts a prosocial and communal orientation toward others with antagonism. It includes traits such as altruism, tender-mindedness, trust, and modesty. Based on this trait the investors are classified as trusting or suspicious investors.
5. Neuroticism

It contrasts emotional stability and even-temperedness with negative emotionality, such as feeling anxious, nervous, sad, and tense. Based on this trait the investors are classified as emotional investors (takes decision based on emotions) or relaxed/confident investors (stable and relaxed in taking decisions).

<table>
<thead>
<tr>
<th>I</th>
<th>Openness to Experience</th>
<th>Vs</th>
<th>Extraversion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Curious investor</td>
<td>Vs</td>
<td>Extrovert</td>
</tr>
<tr>
<td></td>
<td>(Prefers for variety and inventiveness)</td>
<td>Vs</td>
<td>(Talkative and Sociable)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Introvert</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Shy and reserved)</td>
</tr>
<tr>
<td>II</td>
<td>Extraversion</td>
<td></td>
<td>Risk</td>
</tr>
<tr>
<td></td>
<td>Extrovert</td>
<td>Vs</td>
<td>Risk avoider</td>
</tr>
<tr>
<td></td>
<td>(Talkative and Sociable)</td>
<td>Vs</td>
<td>(Prefers to avoid risk)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Risk taker</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Prefers to take risk)</td>
</tr>
<tr>
<td>III</td>
<td>Conscientiousness</td>
<td></td>
<td>Return</td>
</tr>
<tr>
<td></td>
<td>Careless / Spontaneous investor</td>
<td>Vs</td>
<td>Economic investor</td>
</tr>
<tr>
<td></td>
<td>(The people are careless and spontaneous)</td>
<td>Vs</td>
<td>(Highly focused on return than risk)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Highly economical investor</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Thinks that return as a part of investment)</td>
</tr>
<tr>
<td></td>
<td>Dependability</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Independent investor</td>
<td>Vs</td>
<td>Dependent investor</td>
</tr>
<tr>
<td></td>
<td>(They are independent and always take decision on their own)</td>
<td>Vs</td>
<td>(They depend others for information)</td>
</tr>
<tr>
<td>IV</td>
<td>Agreeableness</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Suspicious investor</td>
<td>Vs</td>
<td>Trusting investor</td>
</tr>
<tr>
<td></td>
<td>(They are always suspicious)</td>
<td>Vs</td>
<td>(They believe others)</td>
</tr>
<tr>
<td>V</td>
<td>Neuroticism</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Relaxed / Confident investor</td>
<td>Vs</td>
<td>Emotional investor</td>
</tr>
<tr>
<td></td>
<td>(Stable and relaxed in taking decisions)</td>
<td>Vs</td>
<td>(Take decision based on emotions)</td>
</tr>
</tbody>
</table>