CHAPTER 3
DATA BASE AND METHODOLOGY

This chapter explains the Research Design, the Methodology and the Research Procedures used in the study. The chapter also explains the procedure of data collection and the relevant analysis, techniques and methods. All the problems and issues that have come across during the research process have also been critically discussed.

3.1 Definitions and Research Questions

The literature reviewed helped us to understand that how consumers perceive Ethics in E-marketing. An insight was drawn to understand the prevalence of ethical issues in E-marketing (Privacy, Spamming, Internet Frauds, Spyware and Cookies) and the influence of related unethical practices thereof on the buying behaviour of the consumers’ shopping online. These related terms have been explained as given below:

3.1.1 Definitions

✓ E-Marketing: E-marketing means using digital technologies to sell goods or services. These technologies, like e-mail and websites, are a valuable complement to traditional marketing methods whatever the size of the company or business model. “E-marketing is achieving marketing objectives through the use of electronic communications technology” (AMA, 2004).

✓ Consumers’ Privacy: Consumers’ privacy can be defined on the basis of two dimensions of control i.e. control of information disclosure and control over unwanted intrusions into the consumer's environment. Disclosure of information pertains to the capturing and storing of consumer information in databases. The two main issues for conducting business online are privacy and security. The invasion of privacy on the Internet is commonly regarded as the unauthorized collection, disclosure, or other use of personal information (Wang, Lee, and Wang 1998).

✓ Spamming: Least satisfying dimension of the growing E-marketing environment is the spam unsolicited email that typically attempts to sell products and services to
Internet users. Spamming is defined as an e-mail message of a commercial nature that has been transported without the explicit permission of the receiver (Krishnamurthy, 2000).

✓ **Cookies:** Cookies are small files placed on a user's computer by a third party entity when that person is browsing websites on the Internet. Such cookies record various information of the user that is later retrieved by the computer that placed them on the user’s site. Wikipedia (2009) defines cookie as a text file embedded in an http file that is exchanged between a server and a web browser running on a client, and retrievable by the server.

✓ **Spyware:** Spyware can be defined as an agent technology or controversial software technology that allows the collection of personal information from the computer linked to the internet. The main purpose of the spyware is to collect information, and send it to the information collector. Wilsker (2002) opined that spyware-software tracks almost all surfing activities, and sends that information to a site that compiles this information with other personal information.

✓ **Internet Frauds:** It refers to the crime in which the perpetrator develops a scheme using one or more elements of the Internet to deprive a person of property or any interest, estate, or right by a false representation of a matter of fact, whether by providing misleading information or by concealment of data. Internet Market fraud is any type of intentional deception that uses the Internet. In the present study various kinds of frauds such as Credit Card Fraud, Identity Fraud, Click Fraud, Financial Fraud and Product Fraud etc. have been discussed.

### 3.1.2 Research Questions

✓ What are the consumer’s perceptions and attitudes towards major ethical issues as well as related unethical practices in E-Marketing (i.e. Privacy, Spamming, Spyware etc.)?
✓ How consumers perceive E-marketing and related ethical issues with regard to their Demographic as well as their Psychographic variables?
✓ What are the causes of the violation of consumer’s privacy?
What are the reasons due to which consumers prefer or do not prefer to go for online shopping?
What percentage of respondents is aware of different types of frauds committed due to breach of privacy on the Internet and what preventive measures are often taken by them?
What percentage of respondents is aware regarding the legal / regulatory framework to monitor ethical issues in E-Marketing?

3.2 Research Methods

A research method is an action plan which helps to get the information from the initial set of questions to be answered and hence providing some set of conclusions regarding these questions. It helps the researcher in the process of collecting, analyzing and interpreting the observations. There are two basic research methods: qualitative and quantitative research.

Qualitative research involve making observations, usually to develop a new hypothesis or contribute to a new theory, whereas quantitative research begins with known theory and usually develops by attempting to provide evidence for or against a pre-specified hypothesis. Qualitative analysis often involves finding patterns, ideas, and categories, whereby quantitative analysis is analyzed according to an existing framework (Patton, 2002).

In most cases, quantitative research is a method adopted to ensure objectivity, generalisability and reliability (Amaratunga et al., 2002). The standardized questionnaire and the statistical methods are the some of the best examples to test predetermined hypotheses. The intensities of the quantitative paradigm are that its methods generate quantifiable, reliable data that are usually generalisable to some larger population. Qualitative research methodologies are designed to provide the researcher with the perspective of target audience members (Amaratunga et al., 2002). Qualitative methods used in social marketing include in-depth interviews and focus groups or observations. These methods are designed to assist researchers interpret the meanings people assign to social phenomena and to elucidate the cognitive processes underlying behaviors (Crabtree and Miller, 1999).

The research strategy could be exploratory, descriptive or explanatory. In the early stages the research design had been exploratory then after getting into the problem it was verified and
quantified by conclusive research. The kind of conclusive research design adopted for the present study was descriptive in nature. In this study both quantitative and qualitative research methods had been adopted but more emphasis was laid on the quantitative research. The qualitative research method was only used with a goal of acquiring insight into the E-marketing and related ethical issues.

3.3 Hypotheses of the Study
To study the causal relation existing between the variables in the present study, the following null hypotheses were formulated in a broader way:

**H04.1:** There is no significant association between the perceptual factors of the respondents regarding E-marketing with their attitude towards the phenomenon of e-marketing as a hole.

**H04.2:** There is no significant association between the perceptual factors of respondents regarding E-marketing with their attitude towards ethical concerns of E-marketing.

**H05:** There exists no significant association between the perceptual factors extracted with regard to different ethical issues of e-marketing and the attitude of respondents towards these ethical issues (privacy, spamming, internet frauds, spyware and cookies).

**H06:** There is no significant variation between the responses of the respondents belonging to two different psychographic groups with regard to the statements pertaining to the different ethical and other related issues in E-marketing.

**H07:** There is no significant association between the awareness levels of the respondents regarding identity fraud, Credit Card/financial information Fraud, Mass-marketing fraud, Spam emails and phishing, Cash withdraw fraud, Money Transfer Fraud, Bill payments fraud and Application fraud committed due to breach of privacy in E-Marketing and their gender as well as age attribute.

**H08:** There exists no significant association between the awareness levels of the respondents regarding Data protection Law, Information Privacy Act, Children’s Online Privacy Protection Act, Information technology Act, CAN-
The above mentioned hypotheses were framed and tested for chapter four, five, six, seven and eight.

3.4 Universe of the Study
The universe of the study comprised of one of the prosperous states of India that is Punjab and the Union Territory of Chandigarh.

3.5 Sampling Design
Sampling is concerned with the choice of a subset of individuals from a statistical population to estimate the characteristics of the whole population. In the present study, the sample of the study was based on judgement sampling (Non-probability sampling technique). The districts selected became the first stage of sampling unit. In the view of time constraints as well as the large population of Internet users in the Punjab and Chandigarh, judgement sampling was used to collect data from individuals who could reasonably interpret the E-marketing and form ethical viewpoint toward different ethical issues in E-marketing. Hence, in the present study only those individuals had been included who were educated and were exposed to E-marketing. The survey had been conducted via email and in person.

3.6 Sample Size
The sample size is an important feature of any empirical study in which the goal is to make inferences about a population from a sample. The proposed sample size for the survey was 600. A total of 640 survey questionnaires were sent. Out of these 598 questionnaires were received. Each of the responses received had been screened for errors, incomplete or absent responses. Attempts had also been taken to contact those respondents who had given incomplete information. However, the questionnaires having more than 25% of the questions left unanswered had been discarded for the data analysis. The mid-point scale of 3 was assigned to the other remaining questionnaires which were having less than 25% of the questions unanswered. After the screening process was carried out, only 568 responses had been considered complete and valid for data analysis. This represents a success rate of 94.66%, which was reckoned to be good in view of time and cost constraints.
3.7 Demographic Profile of the Respondents

From a marketing standpoint it is important to profile the online as well as offline buyers before we begin discussing the results of the data analysis, as buyer demographics is the foundation of the market division. This may affect business strategy decisions. The demographic features of different respondents are exhibited in the table 3.1.

Table 3.1

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description &amp; Statistics (No. of respondents and percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>293 Males (51.6%); 275 Females (48.4%)</td>
</tr>
<tr>
<td>Place of residence</td>
<td>373 Punjab (65.7%); 183 Chandigarh (32.2%); 12 Other (2.1%)</td>
</tr>
<tr>
<td>Occupation</td>
<td>154 Business (27.1%); 210 Service (37%); 59 Housewife (10.4%); 138 Student (24.3%); 7 other (1.2%)</td>
</tr>
<tr>
<td>Age</td>
<td>29 Below 15 years (5.1%); 361 15-30 years (63.6%); 93 30-45 years (16.4%); 85 Above 45 years (15%)</td>
</tr>
<tr>
<td>Education</td>
<td>11 Secondary (1.9%); 102 High Secondary (18%); 260 Graduation (45.8%); 195 Post Graduation (34.3%)</td>
</tr>
<tr>
<td>Geographical Location</td>
<td>208 Rural (36.6%); 360 Urban (63.4%)</td>
</tr>
<tr>
<td>Income</td>
<td>94 Below 10,000 (16.5%); 197 10,000-25,000 (34.7%); 108 25,000-50,000 (19%); 64 50,000-75,000 (11.3%); 105 Above 75,000 (18.5%)</td>
</tr>
<tr>
<td>Marital Status</td>
<td>221 Married (38.9%); 347 Single (61.1%)</td>
</tr>
</tbody>
</table>

3.8 Demarcation of the Study

Since the aspects of present chosen study were many, we tried to narrow down the focus. The aim of the research was to find perceptions & attitudes of respondents towards major ethical issues in E-Marketing i.e. Privacy, spamming, internet fraud, spyware and cookies, which had been taken based on the past studies. By reviewing the relevant literature it was noted that definitions of electronic marketing vary according to each author's point of view, background and specialization. In the study the terms E-Marketing, Internet marketing and online marketing were used interchangeably and synonymously as it was borne out by many authors i.e. Smith and Chaffey (2001), Strauss and Frost (2010). So in this study E-marketing was only concerned with marketing through internet.
3.9 Sensitivity

The sensitivity of the scale is an important measurement concept, especially when changes in attitudes are under investigation. Sensitivity refers to an instrument’s ability to accurately evaluate the variability existing in the concept. A dichotomous response category such as disagreeing and agree does allow the transcription of any attitude changes. In this study by adding strongly disagree, disagree, not sure, agree and strongly agree categories sensitivity of the scale used was increased in the subject field.

3.10 Scale Development and Testing

Scale testing is an important prerequisite for data analysis. The use of reliability and validity are common in quantitative research and now it has been reconsidered in the qualitative research paradigm. In the present study following considerations were made for Unidimensionality, Reliability and Validity:

3.10.1 Unidimensionality: It is basically concerned with the existence of a single concept underlying a group of measures. Moreover, it must be conducted before Structural model testing is done (Gerbing and Anderson, 1988). Assessments are based on the Eigen greater than 1.0 for the first dimension and Eigen values less than 1.0 for second supports the constructs exhibit unidimensionality. So each variable must load on a single factor.

3.10.2 Reliability: It is concerned with consistency among multiple measures of variability. Reliability represents the systematic variance of the constructs (Olery and Vokurka, 1998). The extent to which results are consistent over time and an accurate representation of the total population under study is referred to as reliability and if the results of a study can be reproduced under a similar methodology, then the research instrument is believed to be reliable (Joppe, 2000).

There are several procedures to compute the reliability. Some of them are: Alternate or Parallel-forms, Test-Retest (Repetition), Split-half technique and Rational Equivalence. Nevertheless, the single method is not perfect so in the present study we had relied on a series of diagnostic criteria to assess internal consistency as hereunder (Hair et al., 2009):
I. The reliability coefficient (Cronbach’s alpha) $\geq 0.6$ general construct and $\geq 0.4$ for broadly defined constructs.

II. Item-to-total correlation (Correlation of the item to summated scale) $\geq 0.5$

III. Inter-item correlation (Correlation among items within a factor) $\geq 0.3$

IV. The Eigen value should be $\geq 1.0$

V. The constructs should have at least loading of two items

VI. The reliability measures derived from a confirmatory factor analysis focus on constructing reliability where: $\text{Construct Reliability} = \frac{\text{Sum of squares of factor loading}}{\text{sum of squares of the factor loading} + \text{sum of error variance terms for constructs}}$. Various fit indices like TLI, CFI and RMR etc. may also be considered for SEM.

VII. Bartlett’s test of sphericity: a test for the presence of correlation among variables. It shows the statistical significance that the correlation matrix has a significant correlation among at least some of the variables. The value is acceptable for $p$ (level of significance) $\leq 0.05$.

The value of Cronbach’s alpha greater than 0.7 is considered reliable (Nunnally, 1978). Vande Venn and Ferry (1980) Opined that the alpha value may be low (0.4) for broadly defined constructs. In the present study, analysis has been performed to retain and delete scale items for developing a refined reliable scale. Inter-item correlation and Cronbach’s alpha used in the study. Inter-item correlation produce extent to which an item is correlated to another point of the set under consideration and items with small inter-item correlation are considered for deletion (Netemeyer et al., 2003).

3.10.3 Validity: Wainer and Braun (1998) referred the validity in quantitative research as “construct validity”. The construct is the initial concept, notion, question or hypothesis that determine which data is to be collected and how it is to be collected. It includes the following:

3.10.3.1 Convergent Validity: It refers how well the item measures relate to one another with respect to common concepts and is exhibited by having significant factor loading ($\geq 0.5$) of the measurement constructs and the high correlation between items ($\geq 0.3$) within a construct (Anderson and Gerbing, 1984).
3.10.3.2 **Content Validity:** This is a subjective measure but unlike face validity we ask whether the content of a measure covers the broad domain of the content. If a researcher wants to measure introversion they have to first decide what constitutes a relevant domain of content for that trait. It refers to the extent to which a measure represents all facets of a given social construct. Validity measures the degree to which items on the research instrument actually relate to the content of the area under investigation. The content validity of the research instrument used in this research pertains to the extent to which it examines various influencing factors on consumer’s trust factors in online purchasing. Hair et al. (1998) suggest that the best way in which to ensure content validity is to subject the instrument to judgmental validation by experts in the field. In this case to validate the contents in the present study, consultations with the Experts, Professionals and practitioners in E-Marketing had been exercised. Their feedback, both positive and negative, had helped to shape the final version of the questionnaire. Once the scale is developed it was tested through Pre-pilot and Pilot Survey.

3.10.3.3 **Construct Validity:** A construct represents a collection of behaviors that are connected in a meaningful way to create an icon or an idea invented for a research purpose. The output obtained from data analysis in the present study was consistent with theoretical logic.

3.10.3.4 **Discriminant validity:** It shows how well an item measure relates to its hypothesized constructs v/s other constructs in the model (Kerlinger, 1973). It is also translated as a scale adequately differentiates itself or does not differentiate between groups that should differ or not differ based on theoretical reasons or previous research. The empirical test here again is the correlation, but this time the summated scale is correlated with a similar but conceptually distinct measure. The correlation is low demonstrating that summated is different from other similar concepts.

3.11 **Research Process**

Research is not concerned with the revision of the facts and building up to date knowledge but also to expose the new facts involved through the process of dynamic changes in the society. In the present research the necessary steps were as follows:
3.11.1 Problem Definition: The collected information from the review of past literature, content analysis and discussions with the professionals and practitioners in the field of Ethics in E-Marketing, defined the research Problem. The discussion had led to the suggestion that a lot of research work going on Electronic marketing but still no importance had been given to the ethics in E-Marketing. So we decided to focus on Ethics in E-marketing based on research gap analysis. The research problem hence can be stated as to study consumer’s perceptions towards different ethical issues in E-marketing. To accomplish this the objectives as enlisted in the Chapter1 were framed.

3.11.2 Literature Survey: The literature survey concerned with the inspection of past research findings, concepts and theories. In this research we had gone through various national and international book, Journals, Magazines as well as the proceeding of the conferences along with the different websites related to E-marketing and concerned different Ethical issues.

3.11.3 Questionnaire Design: In consultations with professional and practitioner in the field of E-marketing and based on literature survey the non-disguised structured questionnaire was designed and send for pre-pilot study. The questionnaires were received back with suggestions and recommendations, as a result of pre-testing it was modified and revised suitably. Later on based on the recommendations and suggestions questions were improved and sent for final survey. A structured questionnaire was used to gather the necessary data. It served as the primary data to answer the research questions and objectives pertaining to Ethics in E-marketing. The survey questionnaire consisted of eight sections, each of which contains questions pertaining to different parts of the study. The questionnaire comprised of eight parts, first two parts pertained to the respondent’s perceptions towards E-Marketing in general and Unethical practices in it. Rest of the five parts dealt with the respondents perception/awareness towards different Ethical issues in E-marketing and regulatory framework of these issues. The last eight section of the questionnaire dealt with the demographic profile of the respondents viz. age, income, gender etc. The analysis of the responses had been done on the basis of respondent’s demographic and psychographic attributes which were derived from different statements pertaining to ethical and related issues in E-marketing.
3.11.4 **Respondents Selection:** In view of time constraints as well as the large population of Internet users in the Punjab and Chandigarh, judgment sampling is used to collect data from individuals who could reasonably interpret the E-marketing and form ethical viewpoint toward issues in E-marketing, hence in the present study the those individuals had been included who were educated and exposed to E-marketing. The survey questionnaires were only emailed to Internet users who agreed to take part in the survey. This measure was taken to avoid complaints from other Internet users and also to increase the responsiveness of respondents for the survey.

3.11.5 **Data Collection:** A total of 640 survey questionnaires were sent, out of which 598 questionnaires received back. Each of the responses received had been screened for errors, incomplete or missing responses. After the screening process was carried out, only 568 responses were considered complete and valid for data analysis. This represented a success rate of 94.66%, which was reckoned to be good in view of time and cost constraints.

3.11.6 **Data Analysis and Interpretation:** In an effort to sharpen the inferences, various statistical tools and tests were used. Statistical tools like percentages, weighted average scores, Pearson’s correlation coefficient, E test, ANOVA and Chi-square were used. Besides these, Factor Analysis, Cluster Analysis, Multiple Discriminant Analysis and Structural Equation Modeling techniques were also applied using the software like SPSS (16.0), AMOS (4.0) and MS Excel (2007).

- **Weighted Average Score:** This technique was used with an objective to find out the importance given by the respondents to the given precautions against Internet frauds committed due to breach of privacy.

- **Pearson’s Correlation Coefficients:** The main purpose to conduct correlation analysis was to measure whether the main variables in the different ethical issues were independent from each other or not. As a rule of thumb, if a correlation coefficient value of \( r \) indicates 0 to .2, there is a weak relationship between the variables. If \( r \) values of .3 to .6, generally considered moderate, and .7 to 1 is strong (Dancey and Reidy, 2007).
• **Chi-Square Test:** This test was used whether the responses of the respondents were having any association with their attributes of gender and age.

• **Factor Analysis:** The analysis summaries a majority of the information in the data set in terms of relatively few new categories, known as factors. This technique was used to study the various perceptual factors regarding E-marketing and related ethical issues namely privacy, spamming, internet frauds, spyware and cookies. The factor scores from factor analysis of statements regarding E-marketing and related ethical issues were retained and used in SEM and discriminant analysis.

• **Cluster Analysis:** The technique of cluster analysis is ideally suitable for segmentation application. A cluster, by definition, is a group of similar objects, which may be a brand or people based on their attributes towards a particular issue. In the present study, K-Means cluster analysis had been used. This procedure attempts to identify relatively homogeneous groups of cases based on selected characteristics, using an algorithm that can handle large numbers of cases. In the present study this technique was used to segment the respondents into two heterogeneous clusters. Each cluster comprising of sets of respondents having homogeneous attitude and perceptions within the same cluster with respect to ethical concern of e-marketing. Further, the clusters derived were used in the discriminant analysis to study the behaviour of the respondents belonging to a particular cluster towards different ethical issues in E-marketing.

• **Structural Equation Modeling:** SEM includes measurement model and path analysis, is an efficient way to find the causal relationships between constructs and their underlying measurement suitability. In the present study this technique was used to find the influences of extracted factor relating to different ethical issues on the development of attitude towards ethical issues in E-marketing.

• **Discriminant Analysis:** The multivariate technique was used to study the reasons for differences between two or more groups of persons with respect to several variables simultaneously. MDA identifies the variables with the greatest difference among the groups and derives a discriminant-weighting coefficient for each variable to reflect
the differences. In the present study, the MDA was used to map the profile of the respondents belonging to two different groups with respect to their psychographic (extracted using factor analysis) and demographic variables.

3.12 Limitations and Further Research Directions of the Study

As this is the first ever research conducted on Punjab and Chandigarh data, it has set the groundwork for further research. Firstly, random sampling techniques are not used in this research; the ability of the collected data to infer the entire population is reduced because only consumers in principal cities of Punjab and UT Chandigarh were sampled. A random sample covering the whole of Punjab and UT Chandigarh should be applied to test whether the results from this study are replicated. Secondly, the sample size was relatively small. A larger sample should be used to implement any future research in this area. The respondents taken were both online and offline customers and the majority of attendees were college-educated, students or employees in service industries, so the samples may be overly concentrated in some particular consumer groups, rendering bias in measuring the consumers online shopping experience and lifestyle. Therefore, the results obtained may reflect only these groups’ characteristics. It is suggested that future research could possibly diversify into more consumer groups so that the consequences can be applied more generally.