Chapter 5

Research Methodology

5.1 Introduction

Chapter 4 discusses the prominent development of the conceptual model and proposes loyalty and perceived value as the important constructs for this dissertation. The antecedents of loyalty and perceived value in luxury hotels across cultures include functional and technical performance along with brand trust and brand affect. The proposed consequences of loyalty and perceived value are price insensitivity, positive word-of-mouth and perceived brand image. The moderating variables in this dissertation include business/leisure and individualism/collectivism.

This chapter includes the discussion of research methodology, including research design, sampling plan, data collection instruments and procedures, operational definitions of research variables, and analytical measurement. The analytical measurement is divided into the statistical procedures of scale validation, scale dimensionality, confirmatory factor analyses and structural equation modeling.

5.2 Research Design

Research design can be broadly classified into exploratory and conclusive (Malhotra, 2004). The objective of exploratory research is to provide insights and understanding of the problem confronting the researcher. Conclusive research is usually more formal and structured than exploratory research. It is based on large, representative samples, and the data obtained are based on quantitative analysis. The objective of the conclusive research is to test specific hypotheses and examine relationships. Moreover, conclusive research can be either descriptive or causal research. The objective of descriptive research is to describe market characteristics or functions. It is characterized by the prior formulation of specific hypotheses. The objective of causal research is to determine cause-and-effect relationships. The main method of this research is experimentation, whereas the major methods of descriptive research are observation and survey.

This dissertation can be classified as descriptive research because it tested the specific hypotheses and examined the relationships among the variables. Descriptive research seem appropriate in this dissertation because it is preplanned, structured and typically based on large representative samples, which allows the flexibility to apply structural equation modeling. Additionally, survey research is used in this dissertation. This method involves a structured questionnaire given to respondents and designed to elicit specific information. The survey research via questionnaire is appropriate in this dissertation for the following reasons:

1. The questionnaire is simple to administer.

2. The data obtained are reliable due to the limited responses to the alternatives stated. The use of fixed-response questions reduces the variability in the results that may be caused by differences in interviewers.

3. Previous analyses such as Chaudhuri and Holbrook (2001); Patterson and Smith (2003) developed in this dissertation employ the survey research. To extend previous studies, it is preferred to use the same methodology. Additionally, it is the most prevalent methodology in exploring service loyalty (Wiersma, 1995).

4. The data obtained from survey research is often accurate. The reason is that the measurement instrument is specifically developed to address the research questions (Slater, 1995).

5. The coding, analysis and interpretation of the data are relatively simple (Malhotra, 2004).

According to the above reasons, survey research is proposed in this dissertation.

5.3 Sampling

5.3.1 Sampling Frame

The population in this dissertation is the number of tourists who stay in luxury hotels in Thailand such as Four Seasons Resort in Chiangmai province, Dusit

Resort at Pattaya, Chonburi province, etc. As mentioned in Chapter 2, the number of five and four star hotels are 96 (39 for five star hotels and 57 for four star hotels), which are registered with the Thai Hotel Association. Additionally, these five and four star hotels provide 30,940 rooms availability (Thai Hotel Association, 2004), the author assumes one customer per room (30,940 customers). The occupancy rate for hotels in Thailand is about 69.4 percent in 2006. Hence, the population covered in this dissertation is approximately 21,472 guests (30,940 multiplied by 0.69). The sample is the subgroup of the population selected for participation in the study. Five star and four star hotels are categorized as luxury hotels (Lau, Akbar, and Fie, 2005). Therefore, the hotel customers or qualified respondents are any individuals who were temporary visitors, staying overnight at the hotel, and could be classified as business or leisure customers.

5.3.2 Sample Size

Structural equation modeling (SEM) is used in this dissertation. In respect to the sample size requirement of SEM several researchers have suggested that sample size ranging between 100 to 200 is recommended to avoid the problem of misspecification (Pritchard, Havitz et al., 1999). However, Hair, Black et al. (2006) conclude that in one approach, the sample size of 200 is proposed to be the critical sample size to test a model. This is consistent with Anderson and Gerbing (1988) that a sample size of 150 or more typically provides the estimated parameters that have small standard errors to be of practical usefulness. A minimum recommended level is five observations for each estimated parameters in structural equation modeling (Anderson and Gerbing, 1988; Bentler, 1993). The sample size in this dissertation is 487 respondents, which exceeds the critical sample size of 200 or above 150 as mentioned above.

5.3.3 Sampling Method Strategy

Two non-probability sampling techniques, judgmental sampling and convenience sampling, are employed in this dissertation. A judgmental sampling technique permitted the selection of respondents based on the judgment of the author. This technique is subjective and its value depends entirely on the researcher's judgment, expertise, and creativity (Malhotra, 2004). The strengths of employing a judgmental sampling technique are low cost, convenience and less time consumed. The reason to employ this technique in this dissertation is the difficulty to recruit and interview respondents in luxury hotels. Most luxury hotels do not allow interviewers to disturb the hotel guests. As mentioned before, the hotel guests are classified into business and leisure customers. Prior to recruiting the guests in luxury hotels, the author consulted with sales managers or human resource managers in hotels such as The Oriental, Intercontinental, Four Seasons, Hotel Plaza Athenee, etc. to select respondents at the appropriate places. Therefore, the researcher screened both types of customers (business and leisure) by using two different criteria. First, for leisure customers to be approached the researcher considered the popular locations for guests in luxury hotels such as the Chao-Phraya river pier (due to the large number of nearby luxury hotels near this location such as The Oriental, Shangri-la hotel, the Peninsula, etc.), the intersection of Rajprasong Road (including several luxury hotels such as Intercontinental, Four Seasons, etc.), Wat Phra Kaew (the national temple and the most visited place for tourists in Bangkok) and so on. Most leisure customers were selected by this method.

For business customers the same methodology was not suitable for recruitment. The reason is that most business customers come to visit for a short period. In addition, it is quite difficult to reach them at the hotel or popular locations. The researcher therefore, requested the sales managers or sale directors in luxury hotels to distribute the questionnaire to their business customers via e-mail or face-to-face (self-administered questionnaire). This methodology was considered appropriate because it ensured that all respondents were really business customers.

After using judgmental sampling to select the locations to recruited customers, convenience sampling was employed for distributing a questionnaire to leisure customers. This technique was felt to be appropriate to this dissertation's objectives because respondents were selected in the right place at the most convenient time. This technique is the least expensive and least time consuming of all sampling techniques. The sampling units are accessible, easy to measure, and cooperative

(Malhotra, 2004). Prior to distributing the questionnaire to leisure respondents, the author and assistants requested that all respondents provide the name of the hotels where they stayed. If hotel names were on the list of Thailand Standard Hotels (2004) or web sites (such as www.wto.com) as luxury hotels they were included. The author and assistants distributed the questionnaire to respondents. Questionnaires were distributed to 520 customers. The response rate was 94 percent, which accounted for 487 completed questionnaires in this dissertation. The proportion of qualified respondents varied in percentage as shown in Table 5.1 as follows:

Table 5.1: Respondents' profile

Type of Customer	Percentage	Age (years)	Percentage
Leisure	70.2	18-24	6.0
Business	29.8	25-34	31.4
Total	100.0	35-44	34.3
Gender	Percentage	45-54	18.5
Male	53.4	55-64	7.6
Female	46.6	Over 65	2.2
	100.0		100.0

5.4 Data Collection Instruments and Procedures

5.4.1 Data Collection Instruments

In this dissertation, a questionnaire was used to collect the data. The questionnaire originally in English was initially distributed to five participants by employing judgmental sampling. Five participants included lecturers in the hospitality area and service providers in luxury hotels. Hence, the original version was revised by these participants. The new version was distributed to those expertise in the hospitality area in Thailand to avoid mistakes. After the suggestions of the participants in Thailand, the final version of the questionnaire was pre-tested on six

American marketing lecturers in the hospitality area to improve and advise on meaning and style. Finally, the revised questionnaire was provided to the respondents in an English version without any translation. There were two reasons for distributing an English questionnaire version to respondents. First, many guests in luxury hotels have received a high of education level according to younger urban customers, high income, and the openness of new technology such as internet, which lead people to understand English language (Hofstede, 1991; Mishra, 1994; Patterson and Smith, 2003). Secondly, in luxury hotels, it was difficult to translate the questionnaire into several languages for customers of various nationalities.

There were several scale category options in the questionnaire. For example, a five-point Likert scale ranging from strongly disagree (1) to strongly agree (5) was employed for several sections, which included functional and technical performances, attitudinal loyalty, price insensitivity, perceived brand image and individualism/collectivism. Furthermore, a seven-point Likert scale ranging from strongly disagree (1) to strongly agree (7) was also employed for brand trust, brand affect, behavioral loyalty and positive word-of-mouth. A ten-point scale ranging from very poor deal, no value at all, not at all worthwhile and extremely poor value (1) to very good deal, very high value, very worthwhile and extremely good value (10) was also used for perceived customer value. Wording for just less than half of the items was randomly reversed to avoid response bias. The questionnaire consisted of two parts. The first part asked the respondents to provide information on their attitudes towards their luxury hotel. The respondent profile (using multiple choice) was provided in the second part of the questionnaire to overcome any unwillingness to answer questions regarding age, income, etc.

5.4.2 Data Collection Procedures

There are four major modes of survey method, which include telephone interviews, personal interviews, mail interviews and electronic interviews (Malhotra, 2004). In this dissertation, the author does not use telephone and mail interviews because this dissertation focuses on business and leisure customers who do not have much time in Thailand. Referring to Mattila (1999), she applied the probability sampling technique, which a trained research assistant observed the behavior of every third customer and recorded the length of the service encounter. After completion of the interaction, this research assistant approached the customer and got the high rate of participation (approximately 75 percent). However, most luxury hotels in Thailand do not allow interviewers to disturb the guests at the hotel. Initially, the author discussed the sampling technique with expertise in luxury hotels such as sales managers, sales directors, and so on in Thailand. They suggested non-probability sampling technique, which are judgmental and convenience sampling as mentioned before. For leisure customers, the researchers considered the popular locations and distributed a questionnaire to them. Respondents were selected from important spots such as Chao-Phraya River Pier, the intersection of Rajprasong Road and Wat Phra Kaew. Then, the respondents are asked to participate. If they agree, they are given the self-administered questionnaire to complete. Most of the leisure customers are obtained by this methodology. For business customers, the author requested sales managers and sales directors in luxury hotels (such as Intercontinental, Four Seasons Bangkok, Imperial Queen's Park Hotel and Hotel Plaza Athenee Bangkok) to provide the questionnaire to their business customers via e-mail or face-to-face (self-administered questionnaire). Total duration of collecting data is two months for both methodologies. Out of the respondents of 478, 345 respondents are classified as leisure and 142 are business customers.

5.5 Operational Definitions of Research Variables

In this section, the author describes scale application of each construct in the proposed model. It starts from the independent variables (functional performance and technical performance, brand trust and brand affect), followed by mediating variables (attitudinal loyalty, behavioral loyalty and perceived value), dependent variables (positive word-of-mouth, price insensitivity and perceived brand image) and moderating variables (business/leisure and individualism/collectivism), respectively.

5.5.1 Independent Variables: Functional and Technical Performance, Brand Trust and Brand Affect

Functional Performance: Generally, in the service sector, the firm must deliver the core service or technical performance with consistency if it wishes to stay in business (Gronroos, 1983; Kasper, Vries et al., 1999). In addition, customers find and appreciate other benefits which relate to processes or how the service is delivered (or functional quality). Such benefits include social motives to build service relationships as well as the additional things that a service provider may perform to enhance perceptions of the long-term loyal customer (Gwinner, Grensler et al., 1998). In this dissertation, functional performance is based on Gronroos (1983), Parasuraman, Zeithaml et al. (1988) and Sharma and Patterson (1999). Furthermore, Patterson, Mandhachitara et al. (2001) examine the performance items subjected to an exploratory factor analysis (varimax rotation) and revealed two clean dimensions labeled technical and functional performance. In their study, functional performance includes 5 items with 5-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = neither agree/nor disagree, 4 = agree and 5 = strongly agree). The Cronbach's alphas (measure for reliability) for the scales range from 0.91 to 0.95 on four service types, including medical, auto servicing, hairdressing and retail banking (Patterson, Mandhachitara et al., 2001). Examples of statements include "My service hotel takes a personal interest in me". "My service hotel treats me as an individual, not just as another customer". The Cronbach's alpha coefficients are above 0.70, which is consistent with Nunnally (1978), who suggested Cronbach's alpha needs to be above 0.70 to exhibit good internal reliability.

Technical Performance: In a study by Sharma and Patterson (1999), the technical quality scale was based on qualitative interviews and contained 4 items to examine personal financial planning services. The Cronbach's alpha was 0.84, or above the acceptable level. Later on, in Patterson, Mandhachitara et al.' (2001) study, technical performance scale is based on Gronroos (1983), Parasuraman, Zeithaml et al. (1988), Ping (1993), and Sharma and Patterson (1999). They include 5 items with 5-point Likert scale and examine in four service types. The Cronbach's alphas of the scales range from 0.85 to 0.92 (Patterson, Mandhachitara et al., 2001). Also in 2001, Patterson and Smith develop a technical performance scale, which is based on extant literature such as Gronroos (1983), Parasuraman, Zeithaml et al. (1988), and Sharma and Patterson (2000) along with qualitative interviews. In their study, technical performance includes only 4 items with 5-point Likert scale instead of 5 items with 5-point scale as in their previous study. The Cronbach's alphas of the scales range from 0.85 to 0.92 for five service industries, which include hairdressing, medical, travel, auto services and banking in Thailand (Patterson and Smith, 2001). According to the above, technical performance in this dissertation is based on the Patterson and Smith (2001) study, which consists of 4 items on a 5-point Likert scale. The reason to select this scale is because the questionnaire is shorter and provides the same outcome (obtains the same Cronbach's alpha coefficients), which is consistent with Burisch's (1984) advice. Hence, this scale also features good internal reliability (above 0.70). Examples of scale items include "As far as I am concerned, my selected service hotel's performance is excellent". "My selected service hotel usually does its jobs from the outset".

Brand Trust: The measure of brand trust in this dissertation is based on Chaudhuri and Holbrook (2001). Brand trust is measured as a four-item index based on a seven-point rating of agreement (1 = very strongly disagree, 2 = strongly disagree, 3 = somewhat disagree, 4 = neither disagree/agree, 5 = somewhat agree, 6 = strongly agree, 7 = very strongly agree). In the Chaudhuri and Holbrook (2001) study, they tested the scale with 30 brand managers for 149 brands in 49 product categories in the United States. The total number of respondents in their study was 4,470 (30 multiplied by 149). The Cronbach's alpha of this scale was 0.81, which shows good internal reliability (above 0.70). Additionally, Chaudhuri and Holbrook (2002) measured brand trust as a three-item index based on seven-point ratings of agreement (1 = very strongly disagree to 7 = very strongly agree). They tested the scale by applying it to 30 respondents for 137 brands in 45 product categories. The Cronbach's alpha is 0.77, which is less than in their previous study. In this dissertation the author uses a four-item index scale based on the Chaudhuri and Holbrook (2001) study. Examples of scale items are "I trust this selected hotel brand". "I rely on this selected hotel brand".

Brand Affect: The measure of brand affect in this dissertation is also based on Chaudhuri and Holbrook (2001). Brand affect is measured as a three-item index based on a seven-point rating of agreement, which is similar to a scale of brand trust. The Cronbach's alpha of this scale is 0.92. Additionally, Chaudhuri and Holbrook (2002) measure brand affect by using the same scale as in their previous study. The Cronbach's alpha is equal to 0.93, which confirms that this scale shows good internal reliability. Therefore, the author applies a three-item index based on a seven-point rating of agreement as in the Chaudhuri and Holbrook (2001, 2002) study. Examples of statements include "I feel good when I use this hotel brand"; "This hotel brand meets my standards".

5.5.2 Mediating Variables: Attitudinal Loyalty, Behavioral Loyalty and Perceived Value.

Attitudinal Loyalty: The attitudinal loyalty scale in this dissertation was developed by Muncy (1983) and Selin, Howard et al. (1988), which reflected the work of Cunningham (1967). The scale includes a four-item index using a five-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = neither agree/nor disagree, 4 = agree and 5 = strongly agree). This four-item scale asks about brand preferences and the consumer's willingness to switch in given situations. This scale was applied for Pritchard, Havitz et al.'s (1999) study for service patrons. The Cronbach's alpha is 0.91 in their study. Examples of scale items include "I consider myself to be a loyal patron of my selected service hotel." "I am committed to this hotel".

Behavioral Loyalty: Originally, this construct included only three items with a seven-point Likert scale, based on Morgan and Hunt (1994) and Crosby, Evans et al. (1990). The scale is tested in Patterson and Smith's study (2003). The Cronbach's alpha ranges from 0.83 - 0.87 for Thailand and 0.87 - 0.91 for Australia for samples in three industries (travel, medical and hairdressing). Rauyruen's (2005) study develops 5 items with a five-point scale and examines the business environment in Australia. The Cronbach's Alpha is 0.94. This dissertation however, focuses on luxury hotels, which is quite different from any earlier studies. Therefore, the original version questionnaire is revised and modified based on pre-testing and recommendations from experts in the marketing and hospitality fields. Four more items were added. Hence, the behavioral loyalty scale in the final version questionnaire consisted of seven items borrowed from Patterson and Smith (2003), Chaudhuri and Holbrook (2001) and Rauyruen (2005). All seven items are measured on a seven-point Likert scale (1= very strongly disagree, 2 = strongly disagree, 3 = somewhat disagree, 4 = neither disagree/nor agree, 5 = somewhat agree, 6 = strongly agree and 7 = very strongly agree). Examples of statements include "I am not looking for another hotel to replace the present one". "The relationship with this hotel is important to me".

Perceived Value: The measure of perceived value is adapted from existing research by Dodds, Monroe et al. (1991) and Grisaffe and Anand (1998). This scale includes a four-item index on a ten-point scale. This four-item scale includes the benefits obtained from the relational exchange given the prices paid, the time spent, and the effort involved in maintaining a relationship with a focal provider. The Cronbach's alpha is a high 0.92 in both two service contexts (retail clothing and airline) in the study of Sirdeshmukh, Singh et al. (2002), which also demonstrates good internal reliability. Examples of scale items include "For the price you pay for this hotel, would you say the price you pay is". "For the time you have during your stay in this hotel, would you say staying in this hotel is".

5.5.3 Dependent Variables: Positive Word-of-Mouth, Price Insensitivity and Perceived Brand Image

Positive Word-of-Mouth: The scale of positive word-of-mouth is adapted from Zeithaml, Berry et al. (1996). The scale in their study (loyalty scale) includes five items with a seven-point likelihood scale (1 = not at all likely and 7 = extremely likely). The Cronbach's alpha is 0.93 to 0.94 across four companies, which are a computer manufacturer, retail chain, automobile insurer and life insurer. However, the five items in their scale are previously derived from two categories on a five item scale, word-of-mouth (three items) and purchase intentions (two items). In this dissertation, the author adopts only three items from the five-item scale in the Zeithaml, Berry et al.'s study (1996), which includes saying positive things about

XYZ to other people, recommending XYZ to someone who seeks your advice and encourage friends and relatives to do business with XYZ. Additionally, the Liu, Furrer et al. study (2001), adopts the three items from a loyalty scale and test on 285 MBA students from several countries including the United States, Korea, Japan, India, and Thailand. Examples of statements are "Say positive things about this hotel to other people". "Recommmend this hotel to someone who seeks your advice". The Cronbach's alpha in their study is 0.81, which is also above the acceptable level (0.70).

Price Insensitivity: Price insensitivity is developed from Zeithaml, Berry et al. (1996). In their study, the price insensitivity scale includes two items with a seven-point likelihood scale (1= not at all likely and 7 = extremely likely). The Cronbach's alpha coefficient in their study is below 0.60, probably due to too few items in this construct. But after revision and modification based on questionnaire testing and recommendation from marketing and hospitality field academics, the author uses only one item with a five-point Likert scale derived from Zeithaml, Berry et al. (1996) (from 1 = strongly disagree to 5 = strongly agree). The single-item scale is "Would you continue to use the service of this hotel if its prices increase somewhat?" The single item scale is used in this dissertation even though single-item scale can not provide a variance due to random errors, specific items, and methodology (Yi, 1990). The reason is that the single-item is appropriate if the object being rated is simple and unambiguous. Rossiter (2002) calls these "doubly concrete" constructs. Furthermore, Bergkvist and Rossiter (2007) tested predictive validity of single-item versus multiple-item measures of two widely studied constructs: attitude toward an advertisement and brand attitude. They conclude that the single-item is recommended for other constructs in marketing that involve basic objects and basic attributes such as perceptions, intention, or satisfaction, even though the reliability of single-item scale cannot be measured.

Perceived Brand Image: Perceived brand image in this dissertation is adopted from Keller (1993) and Lassar, Mittal et al. (1995). Lassar, Mittal et al. (1995) included social image as a dimension of customer-based brand equity. An image construct has 4 items with a 7-point scale testing with consumer product. The Cronbach's alpha is equal to 0.77, which shows good internal reliability. However, this dissertation applies five items with a five-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = neither agree/nor disagree, 4 = agree and 5 = strongly agree), tested in the Scotch whisky market in Thailand by Mandhachitara (2003). The Cronbach's alpha is equal to 0.77, which is the same as the result of Lassar, Mittal et al. in 1995. Mandhachitara (2003) was adopted due to these items being in a similar luxury product group. Examples of item scales include "This hotel brand has an attractive package for me". "This hotel brand is prestigious and suitable for me".

5.5.4 Moderating Variables: Collectivism/Individualism and Business/Leisure

Collectivism/Individualism: The Collectivism/Individualism measurement is developed from Donthu and Yoo (1998), which is called CVSCALE. This scale includes six items with a five-point Likert scale, which reflects the cultural value of collectivism/individualism, while Hofstede's cultural dimension (1980) is originally operationalized to measure work-related values. The application of CVSCALE is found in several marketing articles such as Donthu and Yoo (1998), Furrer, Liu et al. (2000) and Yoo, Donthu et al. (2001). Additionally, this scale was tested by Anuwichanont (2003) on the service sector in Thailand, with Cronbach's alpha of 0.84, which shows good internal reliability. Examples of statements are "Individuals should sacrifice self-interest for the group (either at school or the workplace)". "Individuals should stick with the group even through difficulties".

Business/Leisure: A binary question was designed to measure this moderator: business/leisure. This question asked respondents the main objective of their trip, of which they will choose either business or leisure. As a nominal scale, Cronbach's alpha was not applicable.

5.6 Analytical Approach

This dissertation uses structural equation modeling (SEM) for the main relationship model. To examine the conceptual model and associated hypotheses in the previous chapter, structural equation modeling is appropriate due to these confirmatory methods (Joreskog, 1978; Bentler, 1983; Browne, 1984) provide researchers with a comprehensive means of assessing and modifying theoretical models (Anderson and Gerbing, 1988). Structural equation modeling requires the following six statistical assumptions to be met (Mueller, 1996):

1. All exogenous and endogenous latent variables have a mean of 0;
2. Structural relations from exogenous to endogenous latent variables are linear;
3. Error terms in equations: a) have a mean of 0 and a constant variance across observations; b) are uncorrelated across observations and c) are uncorrelated with exogenous latent variables.
4. The observed variable covariance matrix is nonsingular.
5. All observed exogenous and endogenous variables have a mean of 0; and
6. Relationships between indicator variables and their associated latent constructs are linear.

This dissertation employs Amos version 4.01 to analyze confirmatory factor analysis in which the maximum likelihood estimation (ML) method is provided. The ML method is used for theory testing and development (suitable for testing our conceptual model and hypotheses), which includes several relative strengths. This method provides the most efficient parameter estimates (Joreskog and Wold, 1982) and an overall test of model fit. Under the assumptions of a multivariate normal distribution of the observed variables, maximum likelihood estimators have the desirable asymptotic, or large-sample, properties of being unbiased, consistent, and efficient (Kmenta, 1971). Moreover, significance testing of the individual parameters and of overall fit is possible.

Achieving unidimensional measurement (Anderson and Gerbing, 1982; Hunter and Gerbing, 1982) is a crucial undertaking in theory and development. A necessary condition for assigning meaning to estimated constructs is that the measures that are posited as alternate indicators of each construct must be acceptably unidimensional. Hence, each set of alternate indicators has only one underlying trait or construct in common (McDonald, 1981; Hattie, 1985). Two criteria representing necessary conditions are used in assessing unidimensionality, which are internal and external consistency. This dissertation measures only the internal consistency of the construct indicators because it is the general approach for assessing reliability. Reliability refers to a scale producing consistent results if repeated measurements are made (Malhotra, 2004). Internal consistency reliability is used to assess the reliability of a summated scale where several items are summed to form a total score. Therefore, coefficient alpha or Cronbach's alpha is the most acceptable measurement for the internal consistency. An alpha value of 0.60 or above for new scales, and 0.70 or above for established scales is considered to be the criterion for achieving internal consistency (Nunnally, 1978; Gounaris and Stathakopoulos, 2004). SPSS version 11.5 was used to calculate the coefficient alpha.

5.6.1 Validity

Validity is defined as the extent to which the differences in observed scale scores reflect true differences among objects on the characteristic being measured, rather than systematic or random error. Maholtra (2004, p.269) stated that "Perfect validity requires no measurement error". Researchers usually assess three types of validity, which are content validity, criterion validity and construct validity.

Content validity or face validity is a subjective but systematic evaluation of how well the content of a scale represents the measurement task at hand. However, content validity alone is not sufficient measure of the validity of a scale. A formal evaluation can be obtained by testing criterion validity and construct validity.

Criterion validity reflects whether a scale performs as expected in relations to other variables selected as meaningful criteria (Malhotra, 2004). Construct validity is the most sophisticated and difficult type of validity to establish. It addresses the question of what construct or characteristic the scale is measuring. It requires a theory of the nature of the construct being measured and how it relates to other constructs. The measurement model in conjunction with the structural model enables a comprehensive, confirmatory assessment of construct validity (Bentler, 1978). The measurement model provides a confirmatory assessment of convergent validity and discriminant validity (Campbell and Fiske, 1959). Given acceptable convergent and discriminant validities, the test of structural model then constitutes a confirmatory assessment of nomological validity (Cronbach and Meehl, 1955; Campbell, 1960). Therefore, construct validity includes convergent, discriminant and nomological validity (Anderson and Gerbing, 1988).

5.6.1.1 Convergent Validity

Convergent validity is the extent to which the scale correlates positively with other measures of the same construct. This validity can be assessed by three different approaches: (1) checking for high correlation among measures of the same construct from the correlation matrix; (2) assessing by factor loadings of indicator variables to ascertain whether they are significantly loaded on their respective constructs, which exceed 0.40 level commonly considered meaningful in factor-analytical investigations (Anderson and Gerbing, 1988) and (3) investigating on the average variance explained (AVE > 0.50) (Fornell and Larcker, 1981).

5.6.1.2 Discriminant Validity

Discriminant validity is the extent to which a measure does not correlate with other constructs from which is supposed to differ. It is assessed by determining whether the correlation between two different measures of the same variable is higher than the correlation between the measure of that variable and those of any other variable (Churchill, 1979) or by determining whether AVE for each construct is higher than the squared correlation between that construct and any other construct (Fornell and Larcker, 1981).

5.6.1.3 Nomological Validity

Nomological validity is the extent to which the scale correlates in a theoretically predicted way with measures of different but related constructs.

5.7 Scale Dimensionality

As mentioned earlier, scale unidimensionality is considered as the prerequisite to reliability and validity. An assessment of unidimesionality should be examined (Anderson and Gerbing, 1988). There is a procedure to consider for checking the scale dimensionality, which is confirmatory analysis (Nunnally, 1978).

Confirmatory factor analysis (CFA) is a theory-testing model as opposed to a theory-generating method like EFA. In CFA, the research starts with a hypothesis prior to the analysis. The hypothesis is based on a strong theoretical and/or empirical foundation. This method, after specifying the 'a priori' factors, seek to optimally match the observed and theoretical factor structures for a given data set in order to determine the 'goodness of fit' of the predetermined factor model (Gounaris and Stathakopoulos, 2004).

According to the above, the author performed only confirmatory factor analysis because all constructs have already been tested by many eminent researchers such as Chaudhuri and Holbrook (2001), Patterson and Smith (2001), etc. Chaudhuri and Holbrook (2001) propose that brand trust and brand affect are two separate constructs. The purpose of confirmatory factor analysis is to test how well the specified measurement model fits the actual data, which is more applicable in this dissertation.

5.8 Structural Equation Modeling

Structural equation modeling (SEM) is a technique that allows a separate relationship for each of a set of dependent variables. SEM provides the appropriate and most efficient estimation technique for a series of separate multiple regression equations estimated simultaneously. It is characterized by two basic components: (1) the structural model and (2) the measurement model. The structural model is the 'path' model, which relates independent to dependent variables. The measurement model allows the researcher to use several variables for a single independent or dependent variable. In this model, the researcher can assess the contribution of each scale item as well as incorporate how well the scale measures the concept into the estimation of the relationships between dependent and independent variables. In this dissertation, the author adopts seven procedures in structural equation modeling (Hair, Black, Babin, Anderson, and Tatham, 2006) as follows:

1. Developing a theory based model. Structural equation modeling is based on causal relationships. Hence, the change of one variable is assumed to result in the change in another variable.

2. Constructing a path diagram of causal relationships. There are two assumptions that apply to a path diagram. First, all causal relationships are indicated. Second, it relates to the nature of the causal relationships that are assumed to be linear. Hence, nonlinear relationships can not be directly estimated in structural equation modeling; however, the modified structural models can approximate nonlinear relationships.

3. Converting the path diagram into a set of structural equations and measurement equations. The objective is to link operational definitions of the constructs to theory for the appropriate empirical test.

4. Choosing the input matrix type and estimating the proposed model. SEM uses only the variance/covariance or correlation matrix as its input data. The measurement model specifies which indicators corresponds to each construct. Then, the latent construct scores are employed in the structural model. As mentioned before, the sample size of 200 is considered to be the 'critical sample size'.

5. Assessing the identification of the model equations. An identification problem is the inability of the proposed model to generate unique estimates. There are four symptoms to detect an identification problem, including very large standard errors for one or more coefficients, inability to convert the information matrix, negative error variances and high correlation (±0.90 or greater) among the estimated coefficients (Hair, Anderson, Tatham, and Black, 1998).

6. Evaluating the results for goodness-of-fit. SEM includes three assumptions as other multivariate methods, which are independent observations, random sampling of respondents, and the linearity of all relationships. After satisfying these assumptions, the offending estimates are examined. The next step is to assess the overall model fit with one or more goodness-of-fit measures. There are three categories for the goodness-of-fit measures, comprising absolute fit measures

followed by incremental fit measures and parsimonious fit measures, respectively. The absolute fit measures assess the overall model fit (both structural and measurement models), with no adjustment for the degree of 'overfitting' that might occur. The incremental fit measures compare the proposed model to another model specified by the researcher. The parsimonious fit measures adjust the measures of fit to provide a comparison between models with differing numbers of estimated coefficients. For the summary, presentation of goodness-of-fit criteria, refer Table 5.2.

Table 5.2: Goodness-of-Fit Criteria:

Goodness-of-Fit Measure	Acceptable Level
Absolute Fit Measures	
1. Likelihood-Ratio Chi-Square Statistic (χ^2)	Non-significance at 0.05 level or 0.10, 0.20 be exceeded before non-significance is confirmed (Hair, Black et al., 2006)
2. Goodness-of-Fit Index (GFI)	GFI > or close to 0.90 and good fit at 0.95 (Hu and Bentler, 1999)
3. Root Mean Square Residual (RMSR or RMR)	RMSR or RMR < 0.08 (Schmacker and Lomax, 1996)
4. Root Mean Square Error of Approximation (RMSEA)	0.05 < RMSEA < 0.08 (Hair, Black et al., 2006)
Incremental Fit Measures	
1. Adjusted Goodness-of-Fit Index (AGFI)	AGFI > 0.80 (Segars and Grover, 1993) or > or close to 0.90 (Hair, Black et al., 2006)
2. Tucker-Lewis Index (TLI or NNFI)	TLI or NNFI > or close to 0.90 (Hair, Black et al., 2006)
3. Normed Fit Index (NFI)	NFI > or close to 0.90 (Schmacker and Lomax, 1996)
4. Comparison Fit Index (CFI)	CFI > or close to 0.90 (Schmacker and Lomax, 1996)
5. Incremental Fit Index (IFI)	IFI > or close to 0.90 (Schmacker and Lomax, 1996)

Parsimonious Fit Measures	
1. Parsimonious Normed Fit Index (PNFI)	A difference of 0.06 to 0.09 are proposed to be indicative of substantial model differences (Anderson and Gerbing, 1982)
2. Normed Chi-Square (CMIN/DF)	χ^2/df in the range of 2 to 1 or 3 to 1 are indicative of acceptance (Carmines and McIver, 1981). In addition, Wheaton, Muthen et al. (1977) suggest that a ratio of approximately 5 or less is reasonable.
3. Akaike Information Criterion (AIC)	AIC close to 0 indicates a very good fit (Hair, Black et al., 2006)

After the overall model fit is examined, the measurement of each construct is assessed for unidimensionality and reliability. Unidimensionality is an assumption underlying the calculation of reliability and is demonstrated when the indicators of a construct have acceptance fit on a single-factor. Beyond examination of the loadings for each indicator, the principal approach used to assess the measurement model is the composite reliability and variance extracted measures for each construct. The composite reliability of a construct is calculated as follows (Fornell and Larcker, 1981):

$$\text{Construct reliability} = (\Sigma\ \text{Std.loading})^2\ /(\Sigma\ \text{Std.loading}\ )^2 + \Sigma\varepsilon_J$$

Where the standardized loadings are obtained directly from the program output, and ε_J is the measurement error for each indicator. The measurement error is 1.0 minus the reliability of the indicator, which is the square of the indicator's standardized loading. The indicator reliabilities should exceed 0.50, which roughly corresponds to a standardized loading of 0.70.

The variance extracted measure is calculated as:

$$\text{Variance extracted} = \Sigma\ \text{Std.loading}^2\ /\ \Sigma\ \text{Std.loading}^2 + \Sigma\varepsilon_J$$

Rule of thumb: The variance extracted value should exceed 0.50 for a construct.

7. Making the indicated modifications to the model if theoretically justified. The proposed model can be modified by examining normalized residuals and modification indices. Residual values greater than ±2.58 are considered to be statistically significant at 0.05 level (Joreskog and Sorboom, 1989). Another indication of proposed model respecifications is the modification indices (MI). MI values greater than 3.84 are considered to be a statistically significant reduction in the Chi-Square (Hair, Black et al., 2006).

5.9 Summary

The purpose of this chapter is to describe the research methodology approaches on which this dissertation is designed and developed. The research design in this dissertation is descriptive research by using survey methodology. Judgmental and convenience samplings are employed from a total sample size of 487 respondents (345 leisure and 142 business travelers). Personal (face-to-face) and electronic interviews are used in this dissertation. Several statistical methodologies are applied, which are validity and reliability measures, scale dimensionality, and structural equation modeling.