CHAPTER THREE

METHODOLOGY AND RESEARCH DESIGN

3.1 Introduction

The main purpose of this study is to propose a model to understand the behaviour of Muslim tourists with a focus on Islamic attributes of destination in the tourism sector. Thus, it is important to determine the methodology that will be applied to achieve the research objective, to explain the way in which the variables will be measured, and present the research design including data analysis techniques. Moreover, the suitable choices of procedures and methods are very important to improve the validity of the study results. Methodology is defined as the nature of research design and methods. It is a research strategy that shows how research is to be conducted. Methods, on the other hand, are instruments employed in the data collection and analysis (Cohen et al., 2003; Sarantakos, 2005).

Thus, in this chapter, the design and methodology of this study is presented. The chapter provides a description of the specific steps that were taken to address the research problem and test each of the study hypotheses. The chapter discusses issues related to the chosen research methodology, as well as the data collection and analysis methods used to conduct the research study. These issues are addressed after taking into account

- The basic research objectives
- The relevant research questions

This chapter first discusses research methodology and then research methods. The rationale of using triangulation methodology as a base is explained, followed by the overall description of the research design, which includes the research process flow chart. Next, are revisited research questions, research objectives, research framework, and research hypotheses. Next, sampling design, data collection methods and analysis techniques to be employed on the data are discussed. Lastly, the chapter summary is presented.

3.2 Research Methodology

Social science can be conducted within a quantitative or qualitative context (Sarantakos, 2005). These approaches were defined by other terminology such as deduction and induction, theory-then-research and research-then-theory or empirical and theoretical (May, 2001).

The theory-then-research (the research empiricists) strategy assumes a hypothesis-testing approach to research. It formulates a hypothesis from theory, and uses collected data to examine them. This includes developing a theoretical model for testing, creating a number of hypotheses that reflect relationships between its components, design research measures to investigate the model, testing the hypotheses using the collected data, and purifying the model and its associated theories (Reynolds, 1979). The key advantage of the positive strategy is that it "allows researchers to test their hypotheses and rely on objective measures (data) to support their findings" (Wicks & Freeman, 1998).

The research-then-theory (the research theorists) strategy considers that empirical studies should not be restricted to make theories get better by testing hypotheses, but should

discover new theories (Strauss & Corbin, 1998). The researchers should begin with observing the phenomenon's aspects and, in turn, look for data to develop theories about them (Reynolds, 1971).

3.3 Research Methods

Research can be categorized generally into two main methods –quantitative and qualitative. Selecting research methods for an exacting research study may be challenging, because decisions cannot be taken on the suitability of a certain method in separation of the context in which the research problem is found (Downey & Ireland, 1979). Both types of research (qualitative and quantitative) are equally legitimate. Hence, when we are talking about the choice of methodology, the question is not about its quality but about its suitability (Sarantakos, 2005). Given the research question(s) under study, a suitable methodology should be selected. Some research questions require certain types of strategies and do not leave much room for flexibility (Remenyi & Williams, 1998).

Quantitative research was initially adopted in the natural sciences. It can be defined as research including the use of structured questions where the response choices have been pre-set and a large number of respondents participated (Creswell, 1994). Quantitative methods are applied within the positivist research pattern using numerical data, and usually, structured and predetermined research questions, conceptual frameworks and designs. It relies on the researcher's capability to assess the investigated phenomena and analyse the collected data statistically. Quantitative methods in business research include questionnaires, and field and laboratory experiments (Cavana et al., 2008).

In contrast, qualitative research is a method that entails collecting, analysing, and understanding data by observing the participants' reactions (Creswell, 1994). Qualitative research is used within the intepretivist philosophy. It is designed to understand the rich, complex and distinctive nature of human phenomena. It includes interviews, focus groups and observations (Cavana et al., 2008). Qualitative research not only utilizes non-numerical and unstructured data, but also, in general, has research questions and methods which are more broad at the beginning, and become more focused when the study proceeds (Punch & Punch, 2005).

The quantitative method focuses more on the generalization of findings statistically, which tries to clarify and predict results by searching for regularities and cause-effect relationship between constructs. Remenyi et al. (1998) pointed out that for quantitative research it is typically obvious what confirmation is required, and this confirmation may regularly be collected within a fixed structure. Thus, in the social sciences, in general, and marketing research in particular, data collection is usually collected by using questionnaires.

In comparison, qualitative methods usually aim at obtaining a richness of detail rather than statistical generalizations. They also aim at detailed descriptions and understanding of the examined phenomenon by observation and involvement (Bryman & Bell, 2007). Punch and Punch (2005) point out that the major distinction of qualitative research is naturalistic, where researchers study people, things and events in their natural settings. Qualitative research and case study fundamentally refer to the same things and are usually considered as identical terms. In this respect, qualitative research is an approach to the study of the social world, which aims to explain and analyse the culture and behaviour of human beings

and their groups from the subjects' perspectives (Bryman & Bell, 2007). However, qualitative research is often criticized for lacking generalizability, being too dependent on the subjective explanations of researchers and being unable to replicate by subsequent researchers (Bryman & Bell, 2007).

Quantitative research is linked with the deductive approach, while qualitative research is linked with the inductive approach (Cavana et al., 2008). Deductive research is the process by which researchers start with theoretical propositions and then move towards solid empirical evidence. However, inductive research is a process by which the researchers examine certain phenomena, which lead to certain conclusions. In this respect, deductive research needs a superior degree of pre-emptive structure in the process of data collection. If the researcher tries to test a hypothesis, the relationship between different variables should be measured (Bryman & Bell, 2007; David & Sutton, 2004). In comparison, inductive methods are exploratory, seeking to develop knowledge of what is going on from the data collected; this does not need predetermined instruments or a method of counting (Bryman & Bell, 2007; David & Sutton, 2004).

Table 3.1 summarizes the differences between quantitative and qualitative methods. It is clearly shown that quantitative research depends on casual relationships and tests hypotheses, while qualitative research can depend on causal or non-causal relationships and meaning is captured and discovered. Quantitative research analyses data statistically in terms of numbers using large samples (cases), while qualitative research analyses data in terms of words from documents using small samples (cases).

Table 3.1

Differences between Quantitative and Qualitative Approaches

Quantitative research	Qualitative research
Reality is objective and singular, and apart from the researcher	Reality is subjective and multiple, as seen by participants in a study
Researcher is independent of that being researched	Researcher interacts with that being researched
Research is assumed to be value-free and unbiased	Research is value-laden and biased, with values generally made explicit
Theory is largely causal and deductive	Theory can be causal or non-causal, and is often inductive
Hypotheses are tested	Meaning is captured and discovered
Concepts are in the form of distinct variables	Concepts are in the form of themes, motifs, generalization
Measures are systematically created before data collection and are standardized	Measures are created in an ad hoc manner and are often specific to the individual setting or researcher
Data are in the form of numbers from precise measurement	Data are in the form of words from documents, observations and transcripts
There are generally many cases or subjects	There are generally few cases or subjects
Procedures are standards, and replication is assumed	Research procedures are particular, and replication is rare
Analysis proceeds by using statistics, tables or charts, and discussing how what they show relates to hypotheses	Analysis by extracting themes or generalization from evidence and organizing data to present a coherent, consistent picture

Adopted from Cavana et al. (2008)

3.4 Triangulation Methodology

The importance of a mixed methodology, involving quantitative as well as qualitative research, is highly applied in this research, as it is believed that both are important in understanding the actual tourism behaviour. In this study, the literature review covers tourism motivation, Islamic attributes of destination, tourist satisfaction, and destination loyalty. As the literature on the Islamic attributes of destination has not been covered enough, resulting in a lack of relevant understanding regarding the practices in the tourism industry, it may be better to start with qualitative research to support the rationale of the results that are to be obtained from the quantitative data analysis.

The researcher finds it necessary to follow triangulation methodology, which is used to create a synergy between the quantitative and qualitative methods. Thus, this two-stage, sequential mixed method study aims at obtaining qualitative results from focus group discussions and in depth interviews with participants in Malaysia, followed up with the statistical quantitative results.

The researcher wants to use the quantitative and qualitative approaches because:

1. This research will develop a construct to measure the Islamic attributes of destination for Muslims. Therefore, the researcher wants to enrich the understanding of this variable in the first stage by using qualitative approaches to explore more items from the participants.
2. According to the research objectives, the researcher will measure the relationships amongst the travel motivation (push and pull), overall tourist satisfaction, Islamic attributes of destination and destination loyalty to build a model. Therefore, the

quantitative approach is suitable for measuring the relationship between these variables.

3. The combinations of using both approaches will enrich the validity of this research, especially because the Islamic attributes of destination are considered a new area of research. Furthermore, using qualitative approach as well as quantitative approach will give more confidence for building such a model.

3.5 Research Design

Yin (2003, p. 20) argues that every type of empirical research has an implicit, if not explicit, research design. In general, the design is the rational sequence that links the empirical data to a study's primary research questions and, finally, to its conclusions. Informally, a research design is a rational plan for going from here to there, here it may be defined as the original set of questions to be answered, and there are answers about these questions. Between "here" and "there" may be found a number of main steps, including collection and analysis of appropriate data. This research was designed to conduct the following steps, as summarized in Figure 3.1, which shows the research process flow chart. The steps include: literature review, research design, data collection, data analysis, and conclusions.

The field research was planned to be conducted sequentially in two stages as mentioned earlier; the first is qualitative and the second is quantitative. The first stage of the field research was planned to consist of in-depth focus group discussions and individual interviews. The FGDs and the interviews are research techniques pertaining to qualitative research (Myers, 2009, p. 121). The aim of the FGD is to collect information through group

interaction, however, interviews are considered an opportunity to discover new information from individuals (Cavana et al., 2008, p. 138).

This combination of FGD and the interview has its advantages. First, identifying a range of Islamic attributes of destination from the FGD. Second, to explore in depth, by conducting interviews using a relatively wide range of participants in order to discover more Islamic attributes and to check the conclusions with the FGD (Morgan, 1996). The advantages of the FGD method of data collection include (Malhotra, 2007, p. 148; Myers, 2009, p. 125):

- Gathering a group of interviewees will generate a broad range of information, insight, and ideas. Thus, the data generated are usually very rich as ideas build and people try to deliver why they feel the way they do.
- Participant's remarks prompt a chain reaction from the other participants.
- Since participants are not asked to answer particular questions, their answers can be impulsive and unusual and should, therefore, give a precise idea of their opinions.
- During the session, the researchers have more control than in participant observation, but less control than in face-to-face interviewing.
- Flexibility of the topic covered.
- Because participants are interviewed simultaneously, data collection and analysis proceed somewhat rapidly.

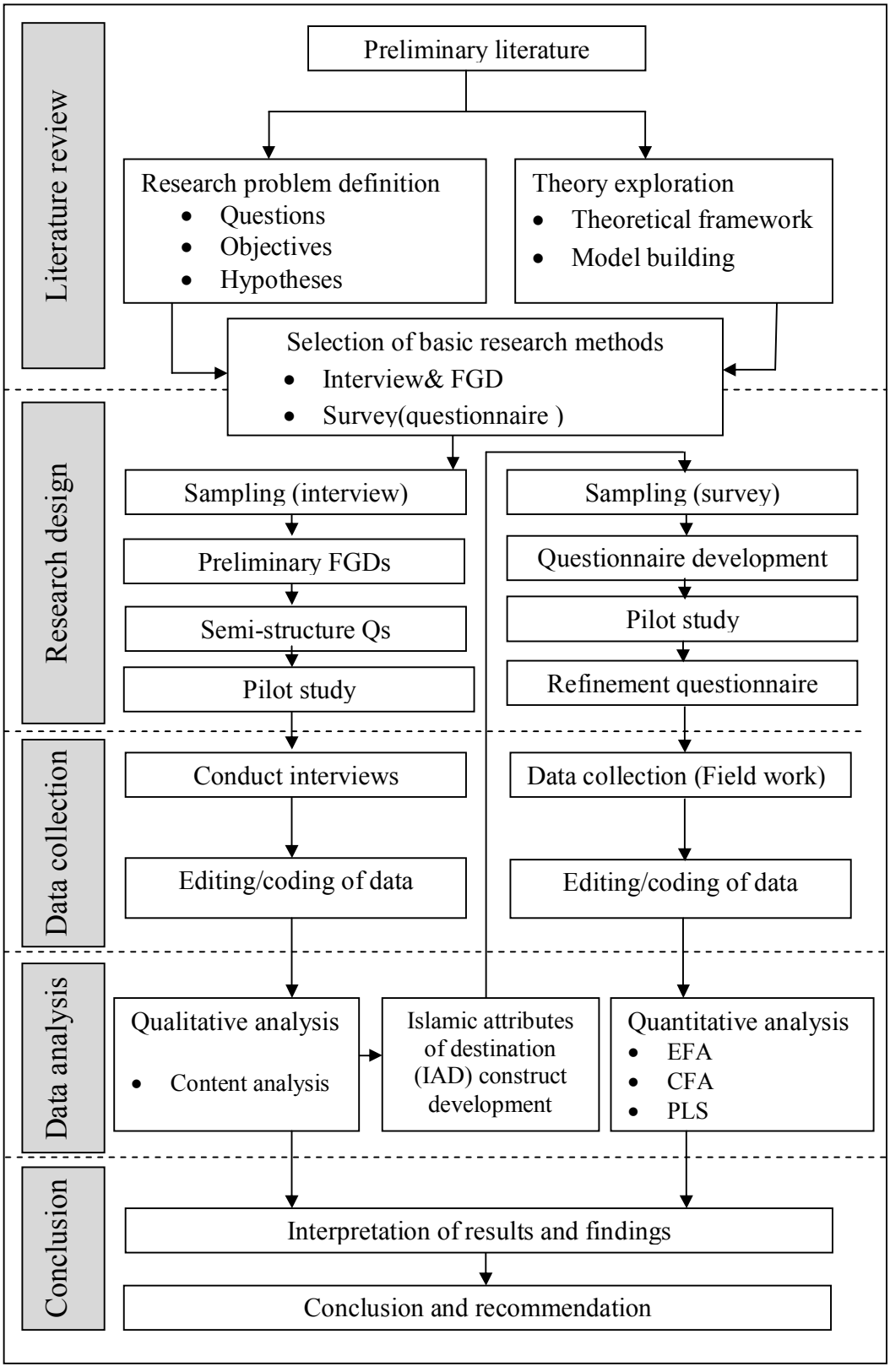


Figure 3.1: The Research Process Flow Chart.

However, there are also some disadvantages of FGDs including (Malhotra, 2007, p. 148; Myers, 2009, p. 125):

- They are often time intensive and costly to run.
- Sometimes difficult to moderate; the quality of the results, thus, depends on the moderator's skills.
- The unstructured comments of participants are difficult to code, analyse, and interpret.
- The results of FGDs are not representative of the population.
- It may be misused by considering the results as conclusive rather than exploratory.

Similar to FGDs, the individual interviews have some disadvantages and advantages (Cavana et al., 2008, p. 151; Malhotra, 2007, p. 155). The advantages include:

- In-depth interviews can expose more depth of insights than focus groups.
- In-depth interviews attribute the answers directly to the participant, unlike focus groups, where it is often difficult to identify which participant made a particular answer.
- In-depth interviews expose a free exchange of information, unlike focus groups, because there is no social pressure to match group answers.
- The researcher can adapt the questions when needed, explain doubts and make sure that the responses are accurately understood by repeating and rephrasing the questions.
- The researcher can also observe non-verbal cues from the participant.

While the disadvantages of in-depth interviews include:

- The participants could feel doubtful of the ambiguity of their responses when they interact directly with the interviewer.
- Sometimes, it is difficult to find skilled interviewers able to conduct in-depth interviews.
- The lack of structure makes the results subject to the interviewer's influence; the quality and comprehensiveness of the results rely much on the interviewer's skills.
- The difficulty of analysing and interpreting the data obtained require skilled psychologists.
- The length and costs of the interview make the number of in-depth interviews small.

Generally, qualitative research is preferable, 'when', 'why', 'who', and 'how' questions are being investigated, when the focus is on an existing phenomenon in reality context, and when the researcher has little control over events (Remenyi et al., 1998). Therefore, the FGD and the interviews provide the basis for this researcher to opt for the survey for the second stage of the research.

The second stage was intended to be a survey research, which is the systematic gathering of data from respondents to understand and/or to predict some aspect of the population of interest. It also implies that the data has been gathered through a questionnaire (Tull & Hawkins, 1987). In the survey research, respondents are required to

answer a range of questions in respect of their behaviour, awareness, attitudes, motivations, intentions, and demographic and lifestyle aspects (Malhotra, 2007, p. 175).

Remenyi et al. (1998) claimed that a survey includes the collection of data from the target population, which is usually known as collecting primary data. It can be used for explanation, description, and/or hypothesis testing; as a guide to action or for analysing the relationships between particular constructs (Burton, 2000). Nevertheless, in social sciences, survey research is typically associated with questionnaires and structured interviews. It can also be integrated with other methods of data collection, for example, in-depth interviews, content analysis, and structured observation (Bryman & Bell, 2007, p. 56).

Survey methods can be categorized according to the mode used to administer the questionnaire. Survey questionnaires may be managed in four major modes; personal interviews, mail interviews, telephone interviews and electronic interviews via e-mail or on the internet. The survey method has several advantages (Malhotra, 2007, p. 175) such as:

- The questionnaire is easy to manage.
- The data gathered are trustworthy because the responses are limited to the alternatives stated.
- The variability in the results that may be caused by differences in interviewers could be reduced because of using fixed-response questions.
- Analysis, coding, and interpretation of data obtained are quite straightforward.

However, there are some disadvantages of survey research, namely (Malhotra, 2007, p. 175):

- Respondents may be incapable or unwilling to give the required information, especially if the information requested is sensitive or private.
- Structured questions and fixed-response alternatives may affect the validity of particular data such as feelings and beliefs.
- Stating the precise words of questions is not simple.
- Question wording may have a major effect on responses.
- Misunderstandings cannot be detected and corrected.
- The seriousness or honesty of responses may not be feasible to check.

Bryman and Bell (2007) point out that surveys are a suitable and helpful method of gathering data under three conditions:

- When the research objectives need quantitative data.
- When the information sought is logically precise and well known to the respondents.
- When the researcher has substantial previous knowledge of certain problems and the variety of responses likely to come out.

However, despite the survey disadvantages, criticisms and conditions, it is by far the most familiar method of primary data collection in marketing research (Malhotra, 2007). However, there is no fully perfect survey, so the key to a successful questionnaire is the care taken in implementing the time consuming preparatory work (Remenyi et al., 1998).

3.6 Research Questions Revisited

The research questions for this study have been mentioned in section 1.6, Chapter One. The main focus of the study questions is on trying to understand Muslim tourist behaviour related to Islamic attributes of destination and its effect on the relationship between travel motivations, overall tourist satisfaction and destination loyalty in the tourism sector of Malaysia. From the main question, there are several questions including: (a)What type of Islamic attributes of destination satisfy Muslim tourists; (b) What push factors of travel motivations are important to Muslim tourists and what is the effect of these factors on overall tourist satisfaction; (c) What pull factors of travel motivations are important to Muslim tourists and what is the effect of these factors on overall tourist satisfaction; (d) To what extent do the Islamic attributes of destination affect the relationship between tourism motivations (pull/push) and tourist satisfaction; and (e) what is the effect of overall tourist satisfaction on the destination loyalty. The answers to question (a) are to come from the interview results (qualitative research) and the answers of questions (b) to (e) are to come from the survey results (quantitative research).

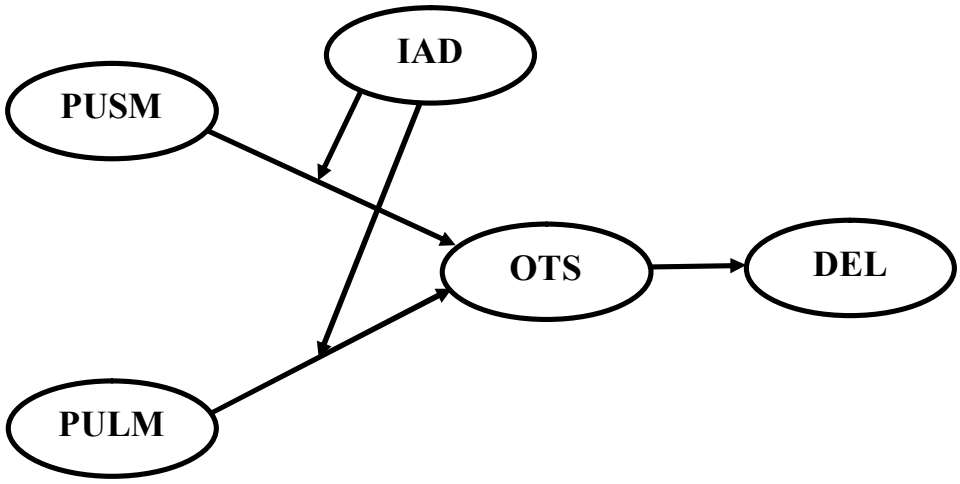
3.7 Research Objectives Revisited

The research objectives of the current study have been mentioned in Chapter One, section 1.6. The main objective of the study is to propose a model to understand Muslim tourist behaviour with a focus on Islamic attributes of destination. There are five specific objectives to assist in achieving the main objective: (a) exploring the Islamic attributes of destination, that satisfy the religious needs of Muslim tourists; (b) determining the important push factors of travel motivations and investigating the possible direct causal effect of these push motivations on Muslim tourist satisfaction; (c) determining the

important pull factors of travel motivation and investigating the possible direct causal effect of these pull motivations on Muslim tourist satisfaction; (d) investigating the moderating effect of the Islamic attributes of destination on the relationship between tourism motivations and tourist satisfaction; and (e) investigating the effect of Muslim tourist satisfaction on the destination loyalty.

3.8 Research Framework Revisited

The framework for this research was developed by relationships assumed from reviewing tourism and destination marketing literature in Chapter Two. The framework for this research proposes that there is a direct link between push motivations (PUSM) and overall tourist satisfaction (OTS). At the same time, the framework shows that a direct link exists between OTS and destination loyalty (DEL). In addition, the framework also proposes that pull motivations (PULM) would have a direct link with overall tourist satisfaction (OTS). In the same framework, a moderate effect has been proposed between PUSM and OTS by Islamic attributes of destination (IAD). Furthermore, the moderating effect also exists between PULM and OTS by IAD. Figure 3.2 presents the study framework. It shows that there are two exogenous variables; push motivations (PUSM) and pull motivations (PULM), two endogenous; overall tourist satisfaction (OTS) and destination loyalty (DEL), and a moderating variable; Islamic attributes of destination (IAD).



Key: PUSM – Push motivations; PULM – Pull motivations; OTS – overall tourist satisfaction; DEL – destination loyalty; IAD – Islamic attributes of destination

Figure 3.2: Research Framework

3.9 Research Hypotheses

It was noticed in the literature review that the influence of travel motivations on overall tourist satisfaction has been studied by various research (e.g. Dunn Ross & Iso-Ahola, 1991; Fang et al., 2008; Fielding et al., 1992; Yoon & Uysal, 2005). Each variable, push motivations (PUSM) and pull motivations (PULM), have hypothesized effects on overall tourist satisfaction (OTS).To test the relationship between travel motivations (pull and push) and overall tourist satisfaction the following hypotheses are proposed (See figure 3.3):

H1: the push motivations (PUSM) positively influence overall tourist satisfaction (OTS).

H2: the pull motivations (PULM) positively influence overall tourist satisfaction (OTS).

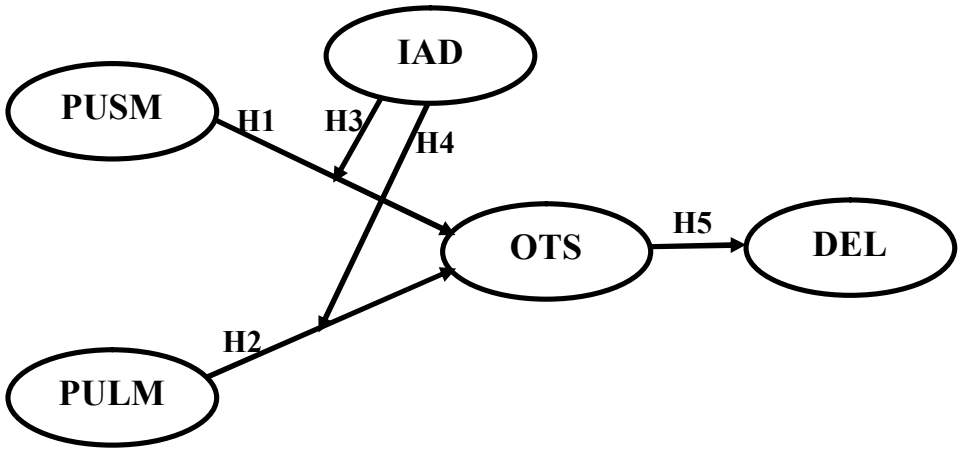
The impact of religion on behaviour and purchasing decisions has been discussed in previous studies (Delener, 1990; Essoo & Dibb, 2004; Hirschman, 1981). Weidenfeld (2006) supports that the availability of religious needs for tourists will increase their satisfaction. Through the framework, it was posited that Islamic attributes of destination (IAD) moderate the relationships between push motivations (PUSM) and overall tourist satisfaction (OTS), as well as the relationship between pull motivations (PULM) and overall tourist satisfaction (OTS). Using the propositions arrived at through the literature review, the following hypotheses were developed:

H3: Islamic attributes of destination (IAD) moderate the relationship between the push motivations (PUSM) and overall tourist satisfaction (OTS).

H4: Islamic attributes of destination (IAD) moderate the relationship between the pull motivations (PULM) and overall tourist satisfaction (OTS).

In the marketing literature, it is acknowledged that satisfaction and loyalty are associated (Oliver, 1999). Destination marketing literature has also noted that overall tourist satisfaction influences the destination loyalty (Akama & Kieti, 2003; Alegre & Cladera, 2006; Baker & Crompton, 2000; Chi & Qu, 2008; Fang et al., 2008; Um et al., 2006; Yoon & Uysal, 2005; Yuksel, 2001). Empirical evidence suggests that overall tourist satisfaction is a strong indicator to suggest the destination to others and to repeat visit (Kozak, 2001; Yoon & Uysal, 2005). Delighted tourists are more willing to share their positive experience of traveling with their relatives and friends, and are more likely to revisit the same destination. To test the relationship between OTS and DEL, the following hypothesis was drawn:

H5: The overall tourist satisfaction (OTS) positively influences destination loyalty (DEL).



Key: PUSM – Push motivations; PULM – Pull motivations; OTS – overall tourist satisfaction; DEL – destination loyalty; IAD – Islamic attributes of destination

Figure 3.3: The Research Hypotheses

3.10 Sampling Design

In the following sub-sections the discussions on sampling design in the two stages of the field research are presented.

3.10.1 Interview Sampling Design

The selection of participants in qualitative research is considered an important task to ensure that participants willingly provide information representative of the target population (Cavana et al., 2008, p. 137). However, Bryman and Bell (2007) argue that the issue of representativeness is less important in qualitative research than in quantitative research because it aims to make an in-depth analysis. Thus, the sample size of this stage was determined by a combination of a lack of extra participants and data saturation. The period of qualitative data collection ranged from June 2009 to October 2009. The start

month was selected because it represents the high season of tourism in Kuala Lumpur, the capital city of Malaysia.

According to Cavana (2008, p. 137), the non-probability methods have a discrete advantage of rapidly accessing participants who are most likely to give rich information. Therefore, the participants included in the focus group interviews and in-depth interviews were to be selected following a convenience sampling approach. The approach basically involves collecting information from the population members who are conveniently available to give it (Cavana et al., 2008, p. 262). Moreover, convenience sampling is most often used during the exploratory phase of a research project and is perhaps the best way of collecting basic information rapidly and efficiently (Cavana et al., 2008, p. 263). For this reason, the convenience approach is considered the best method, as the purpose of the stage is to explore more items relating to the Islamic attributes of destination.

The optimum size of focus group interview is considered an issue for debate. For example, Malhotra (2007, p. 140) reported that groups of less than eight are not likely to generate the momentum and group dynamics required for a successful session. He also reported that groups of more than twelve participants may be too crowded and may not be conductive to a cohesive and natural discussion. Cavana et al. (2008, p. 155) argued that small groups are likely to work best when the participants are interested in the topic while large groups require considerable skill to manage the group discussion and tend to become more complex and formally structured. Thus, Malhotra (2007, p. 140) suggests a rule of thumb range of eight to twelve members, while Cavana et al. (2008, p. 155) suggest six to nine.

It was planned to conduct FGDs first followed by individual interviews. The knowledge gained through the FGDs was then applied to the next stage, which involved preparing questions for semi-structured interviews. Semi-structured interviews were selected as a research method to support the survey design because the objective of this stage is to generate more IAD items. The semi-structured interviews involve the use of some pre-formulated questions, however, there is no strict adherence to them as new questions might arise during the conversation (Myers, 2009, p. 124). Thus, it gives the interviewee the opportunity to add important insights.

3.10.2 Survey Sampling Design

The required sample size relies on factors such as the proposed data analysis techniques (Malhotra, 2007). One of the proposed data analysis techniques for this research is CFA, which is very sensitive to sample size and less steady when estimated from small samples (Tabachnick & Fidell, 2007, p. 682). The literature review indicated that there are no generally accepted criteria for determining a precise sample size using CFA, which is a part of SEM (Hair et al., 2010). However, general guidelines have been proposed.

Hair et al. (2010) suggests that the minimum sample size is 100 when considering models containing five or fewer constructs, each with more than three items with high item communalities (0.6 or higher); 150 when models contain seven or fewer constructs and modest communalities (0.5); 300 when models contain seven or fewer constructs and low communalities (0.45), and/or multiple under identified (fewer than three items) constructs; and 500 when models contain a large number of constructs, some with lower communalities, and/or having fewer than three measured items. This suggestion provides a

rough idea about the sample size for this study. It is generally regarded that 100 is the practical minimum size for using SEM (Hair et al., 2010). Also, Cliff (1987) recommends a sample size of 150 for 40 variables (item statements) in a scale.

A priori statistical power analysis can be calculated using various heuristics. Used as a long-time standard for structural equation modelling using AMOS, sample size should be at least five times the number of indicators (Pallant, 2005). Since the proposed model has five variables with two exogenous variables involving fifty-four indicators; 30 for PUSM and 24 for PULM, two endogenous variables involving eight indicators; 4 for OTS and 4 for DEL, and one moderating variable (IAD) involving 18 indicators. Thus, there are 80 indicators (80 x 5 = 400), which mean that 400 is the required sample size.

An alternative method states that when using the technique of partial least square (PLS) , as in this study, the sample size is independent of the number of indicators when the model is reflective (Chin et al., 2003). The heuristic requires ten times the construct with the largest number of structural paths, which would be either of the two exogenous variables. This method indicates 10 x 30= 300 as an adequate sample size. As this research utilizes PLS as well as CFA, the first heuristic will be considered adequate. The usable sample size of 508 exceeds the sample size of 400, which is deemed adequate by the power calculations.

3.11 Data Collection

There are two stages in this research to collect data during the fieldwork –interview data collection, which will be applied by conducting focus group discussions and individual

interviews, and survey data collection by distributing questionnaires. In the following sub-sections the discussions on tools and procedures for data collection in the two stages of the field research are presented.

3.11.1 Interview Data Collection

It was planned to conduct two focus groups discussion (FGDs) for identifying a range of Islamic attributes of destination, which could further help in conducting the individual in-depth interviews. In response to that, two focus group discussions were conducted in June 2009. The participants of these focus groups were international PhD students currently studying in Malaysia in different areas of research. The method used was the 'convenience sampling method', as mentioned earlier in this chapter and this was carried out by contacting students in the PhD programme. All of them are Muslims.

The participants of FGD 1 consisted of eight males from Saudi Arabia, Palestine, Mauritania, Oman, Jordan, Yemen, Egypt and Sudan. The duration of FGD 1 was one and a half hours. The participants of FGD 2 consisted of seven females from Egypt, Libya, Yemen, Nigeria, Oman, Turkey, and Algeria. The duration of FGD 2 was one hour and fifteen minutes and the participants did not agree for the session to be digitally recorded. The reason for this may be culturally related. All FGD participants had overseas travel experience for various purposes such as tourism, business, learning, and visiting friends.

Table 3.2 shows the interview guideline questions designed according to the knowledge gained from the two FGDs. A pilot study was used on four friends to determine if the semi-structured questions are easily understood and rational. Moreover, the pilot

study was to establish if the questions were too wide or narrow and to practice illustrative questions and probing. In other words, it tested the validity of the proposed questions.

Table 3.2: Interview Semi Structured Questions

	Interview guideline questions
Push motivations	1. What are the reasons that motivate you to travel for tourism?
Pull motivations	1. What are the attractions that influence your decision to travel to a specific destination? 2. Which attributes do you like to be in the destination you will visit?
Islamic attributes	1. In your opinion, which Islamic attributes in the destination do you need when you travel for tourism? 2. If you are planning to visit a Muslim country, which Islamic attributes or aspects of this country are important to you as a Muslim? (Probing; facilities, food, places of worship, environment, behaviour, entertainment, dress, etc.)
Worship	Can you tell me in detail what worship facilities in your travel are important to you?
Hotels	Can you tell me in detail what Islamic aspects of hotels in your travel are important to you? (Probing; food, prayer facilities, entertainment, etc.)
Transportation	What Islamic aspects of transportation to/in the destination are important to you?
General environment	What Islamic aspects of the general environment if you visit a Muslim country are important to you? (Probing; alcohol, gambling, behaviour, dress code, prostitution, etc.)
Religious motives	What are the most likely Islamic attributes that influence your decision to choose the destinations you will visit?

The participants were asked questions relating to why they travel for tourism and select specific destinations. Moreover, they were asked about the Islamic attributes they would require when traveling in a Muslim country or a non-Muslim country. Specifically, the participants were asked about these attributes when visiting Muslim countries. This was followed by probing questions that sought to explore in detail specific attributes generated from FGDs such as access to a mosque, prayer facilities/room at tourism sites, presence of loud public pronouncement of Azan, placement of Qibla stickers, provision of a copy of the Holy Qur'an in hotel room, Halal issues (food, alcoholic drinks, kitchen, gambling), segregated services provided in hotels, Islamic dress code, prostitution, display of affection between sexes and general morality.

A total of 53 interviews were conducted with tourists in Kuala Lumpur, the capital city of Malaysia; 27 interviews in July 2009 and 26 interviews in October 2009. The interviewees were patrons of international hotels in Kuala Lumpur around the 'Golden Triangle Area' where many international chain hotels are located. Two researchers approached as many Muslim tourists as possible at these hotels and requested them to participate in the study.

Thirty-eight tourists did not agree to be interviewed. The duration of most interviews was between 35 to 45 minutes; 41 interviews were recorded digitally. The rest were unable to be recorded due to objections from the interviewees who were females from the Middle East. The interviewer transcribed both the FGDs as well as all interview sessions. Some interviewees from the Middle East preferred to communicate in Arabic while the rest answered in English. The interviewer adopted the probing technique during the session to encourage the participants to provide more details.

All participants (FGDs and interviewees) were asked to describe themselves whether they are secular, mildly religious, religious or very religious. Some demographic information was also gathered. The digital recording of the interviews was changed into text. This study was targeted to discontinue conducting in-depth interviews when no additional information was provided by the participants. The items generated as a result of the interviews provided the opportunity to develop the measures of the IAD construct.

3.11.2 Survey Data Collection

A structured questionnaire was used as the main tool to collect data in the second stage of field research for this study. The study applied the encouragement technique to increase the response rate by giving rewards such as a pen, souvenir, and so forth. The study aimed to reduce the failure of respondents to answer some items in the questionnaire through the development of clear instructions on how the questions had to be answered. In addition, the design of the questionnaire was attractive to encourage the respondents to complete it, for example: cover with nice colour, high quality printing, only six pages, and no crowded sentences.

The questionnaire consists of five parts. Part one includes two sub-sections; section one covers push motivations (PUSM), and section two covers pull motivations (PULM). Part two to part four covers Islamic attributes of destination (IAD), overall tourist satisfaction (OTS), and destination loyalty (DEL). Part five includes general information questions. The measurement items of four constructs were established in literature by many empirical studies. Table 3.3 shows the source of these items as well as newly developed variables. Part one contains 30 push motivations items, and 24 pull motivation items. Part

two contains 18 Islamic attributes items. Part three includes 4 overall satisfaction items. Part four includes 4 destination loyalty items. Lastly, part five includes 8 general information questions such as gender and income, and open-ended questions.

Table 3.3: Sources of Measurement Items

Variable	Sources of measurement items
Push motivation (PUSM)	(Baloglu & Uysal, 1996; Jang & Cai, 2002; Kim et al., 2007; Yoon & Uysal, 2005)
Pull motivation (PULM)	(Baloglu & Uysal, 1996; Jang & Cai, 2002; Jang & Wu, 2006; Yoon & Uysal, 2005)
Islamic attributes of destination (IAD)	Newly developed
Overall Tourist satisfaction(OTS)	(Bigné et al., 2005; Chi & Qu, 2008; Del Bosque & Martín, 2008; Yoon & Uysal, 2005)
Destination loyalty (DEL)	(Bigné et al., 2005; Chi & Qu, 2008; Del Bosque & Martín, 2008; Yoon & Uysal, 2005)

The items statements in part 1 to part 4 are measured as subjective estimates using a five point Likert scale. In the first two parts (PUSM, PULM, IAD), 1 indicates 'not at all important' and 5 indicates 'very important', whereas for the other two parts (OTS, DEL), the five point Likert scale includes the items statements relevant to each variable from 1 to 5. The questionnaires were distributed in two versions –Arabic version and English version. To obtain the equivalent version in Arabic, the process of "back translation" was followed. The study used a questionnaire in Arabic because some of the tourists visiting Malaysia from Arab countries may not understand English. The full set of the English version is found in Appendix A, and the Arabic one is found in Appendix B.

Survey Questionnaire Validity

This research aimed to establish the questionnaire validity by checking content validity, also known as face validity, which refers to the extent that the measure apparently

reflects the content of concept in question (Bryman & Bell, 2007, p. 165). The objective of checking content validity is to ensure that the selection of scale items extends past just empirical issues to also include theoretical and practical considerations (Hair et al., 2010, p. 125). Content validity is usually established by asking other people whether or not the measure captures the concept that is the focus of attention, and is pre-tested with multiple subpopulations (Bryman & Bell, 2007; Hair et al., 2010).

The items used in this study were created based on an extensive review of tourism and destination marketing literature. Some of the studies reviewed are listed in Table 3.3. To establish the content validity of the survey instrument, as suggested by Bryman and Bell (2007), and Hair et al. (2010), five academicians were asked to review the questionnaire to determine whether the measures really captured the concept or not, and to check the correctness of the wording used. Moreover, the questionnaire was also pre-tested on some of the participants to make sure questions were clear and understandable.

Based on the suggestions of the academicians and the pilot study results, any items perceived as ambiguous were either modified, or eliminated. Seven items were removed: one item from the push motivation section (No.1) and six items from the pull motivation section. The items removed were: (1) Meeting people of opposite sex (push motivation), (2) Casino and gambling, (3) Live theatres/concerts, (4) Night life and entertainment, (5) Local cuisine, tennis, and (6) Inexpensive restaurants.

After the content validity of the questionnaire was established, the questionnaires were distributed (Administered from February to May 2010) in international hotels and tourism sites in four Malaysian cities, namely, Kuala Lumpur, Kula Terengganu, Penang,

and Johor Bahru. The questionnaires were distributed in these cities because of the availability of many hotels and tourism sites, and because most of the tourists visiting Malaysia take tours of these cities. The selected hotels were three, four, and five stars. By taking the permission from the hotel, the questionnaires were distributed to tourists and they were informed to leave the filled questionnaire with hotel reception counter. The tourism sites were selected according to the famous sites for each city which normally the tourists visit such as shopping mall, parks, resorts, etc. the questionnaire were collected directly after the tourist filled up the questionnaire in tourism site.

3.12 Data Analysis

Both qualitative and quantitative data collected were analysed. The data analysis techniques to be used in this research for both stages of field research data are discussed in the following sub-sections.

3.12.1 Interview Data Analysis

The purpose of conducting semi-structured interviews, as mentioned earlier, is to explore the Islamic attributes of destination. The research was based on sequential procedures, where the researcher aims to expand the findings from the focus group interviews to the face-to-face interviews. Therefore, two focus group discussions (FGDs) were conducted followed by 53 individual interviews. The intention of this structure is to begin with focus groups to collect initial information to be used in modifying subsequent interview questions based on new knowledge. Thus, the plan for the analysis of the interviews was to divide the analysis into two stages; the first stage was to deal with the FGDs and the second with face-to-face interviews.

The content analysis was used to analyse the interviews (Myers, 2009, p. 172). The constant comparative method describes the process of conducting content analysis (Cavana et al., 2008, p. 171). The researcher reviewed the raw data and identified a critical point and continued reviewing until a distinct second theme emerged. The researcher contrasted this second theme with the first to ascertain that the two are unique and then continued reading until an obvious third theme was found. The researcher then compared this third theme with the first and second themes. This process of constant comparative analysis proceeded, with each newly identified theme weighed against earlier identified themes to ensure that the new theme did shed more light on the phenomenon being investigated.

This process involved coding the participants' comments with a concise summary statement that delineated the Islamic attributes of destination. Two separate summary statements were written for some interviews where two dissimilar points arose in the same comment. The second step entailed separating the statements into different categories. These initial categories were then analysed for rule formation to identify inclusions and exclusions. Categories were then weighed against one another to clarify groupings, merge similar categories, and reject redundancies.

3.12.2 Procedures of Developing IAD Construct

This study followed the sequence of steps that should be performed in developing measures of constructs, as suggested by Churchill (1979), Malhotra (2007, p. 274) and Chen and Paulraj (2004). Figure 3.4 shows the suggested procedure for developing IAD measures. In the following sub-sections these steps are discussed.

3.12.2.1 Specify domain of the construct

The suggested procedure for scale developing begins with specifying the domain of the construct and underlying the theory of the construct being measured. A theory is essential not only for constructing the scale but also for interpreting the resulting score. Therefore, what is included in the construct and what is excluded should initially be investigated. In this stage, determining the domain is applied by conducting a thorough review of literature in which the variable is used and should present a detailed statement of the reasons and evidence.

3.12.2.2 Generate sample of items

The second step in the procedure for developing better measures is to generate items that capture the domain as specified. Typically, this can be based on exploratory research; including literature searches, interviews, and focus groups. The literature should specify how the variable has been defined previously and how many dimensions or components it has. The individual interviews and focus group discussions can also be used in favour of the item-generation stage.

Early stages of item generation focus on developing a set of items that capture each and every one of the dimensions of the construct in this matter. Near the end of the release stage of item development, the focus will shift to editing the item. Each statement should be reviewed so that the formulated words would be as accurate as possible. After the gathered item is edited carefully, the items generated from both interviews and past literature should be submitted to a panel of experts and/or other knowledgeable individuals for judgment. The panel of experts may refine the items and some items may be dropped. Further refinement will also be applied in the actual data.

3.12.2.3 Purify the measures

Although content validity was achieved in the previous stage, the reduced set of items may still be too large to represent a scale. Thus, further reduction techniques are applied in a quantitative manner. Data should be collected on the reduced set of potential scale items from a pre-test sample of respondents. The data were analysed using techniques such as coefficient alpha and factor analysis. The Cronbach's alpha value is usually used in the early stage to purify the measures. The component is accepted if the Cronbach's alpha value is greater than the .70 threshold recommended by Nunnally and Bernstein (1994). In order to gain the highest possible reliability coefficient, the components are purified by dropping items with the lowest item-to-total correlation. This study used 60 true questionnaires in this stage. Once a satisfactory Cronbach's alpha coefficient is achieved, the analysis moves on to the next stage of instrument development.

3.12.2.4 Continuous improvement cycle

The continuous improvement cycle in the instrument development process was suggested by Chen and Paulraj (2004). It is similar to the process recommended by Churchill (1979) covering: purify measure by exploratory factor analysis, assess the reliability with new data, and assess construct validity. Table 3.4 gives guidelines on the tests to be used in achieving the continuous improvement cycle process. By using 153 true questionnaires, the EFA approach was applied first to establish construct validity, and then the CFA approach, which is discussed in detail in the next section. This is followed by analysis of the survey data. Once the internal consistency and construct validity are satisfactory, that means the measures are valid and reliable. In other words, the reliability

and validity of IAD measurement scale are determined and the IAD instruments can be used for further analysis.

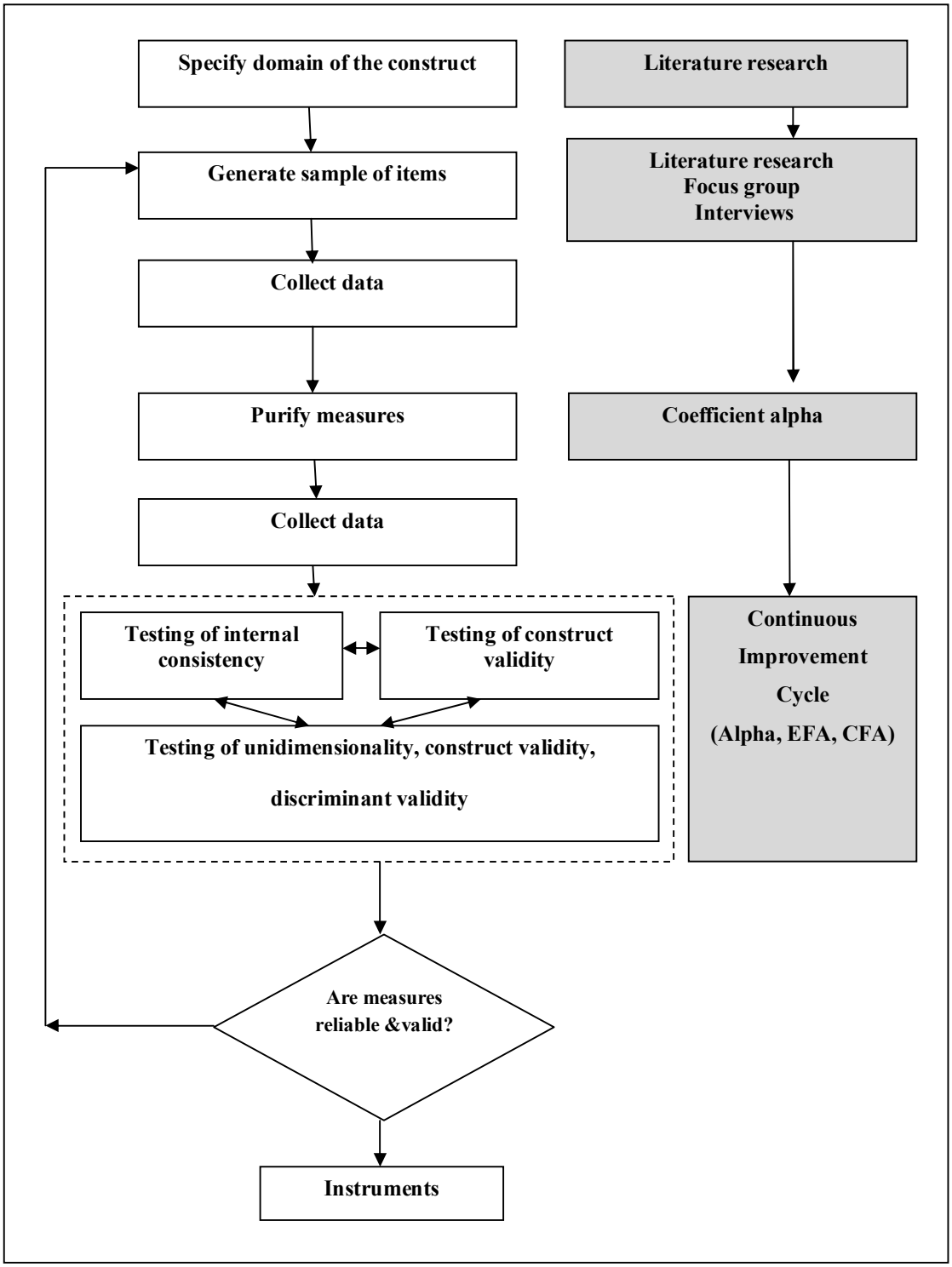


Figure 3.4: Suggested Procedure for Developing IAD Measures

Table 3.4: Tests for the Continuous Improvement Cycle

Test	Procedure
Internal consistency	Cronbach's alpha
Construct validity (EFA approach)	Unidimensionality: Factor loadings Convergent Validity: Eigen value, Variance Extracted-VE, Reliability
Construct validity (CFA approach)	Convergent Validity: *t*-values, squared correlations Fits and Unidimensionality Assessment: Fits and indices Discriminant Validity: constrained and unconstrained model pairs;Variance Extracted versus squared correlation between factors Composite Reliability; Variance Extracted

3.12.3 Analysis of Survey Data

The analysis of the survey data was conducted by applying statistical techniques. These techniques involve descriptive statistics, assumptions for data analysis (e.g., normality, linearity), quantitative data analysis using EFA, CFA, and PLS. Confirmatory factor analysis (CFA) was used to validate construct measurements. PLS was used for assessing the hypothesized relationships contained in the hypothesized model. These are described in the following sub-sections.

3.12.3.1 Procedures for exploratory factor analysis

Exploratory factor analysis (EFA) seeks to discover the underlying structure of a relatively large set of variables. Any indicator may be associated with any factor, and this is the most familiar form of factor analysis. In other words, factor analysis tries to identify underlying variables, or factors that explain the pattern of correlations within a set of observed variables.

Factor analysis is frequently used in data reduction to categorize a small number of factors that explain most of the variance observed in a much larger number of manifest variables (Hair et al., 2010, p. 94). However, factor analysis may be used for any of the following reasons (Gorsuch, 1983):

1. To decrease a large number of items to a smaller number of factors for modelling purposes, where the large number of items prevents modelling all the measures individually.
2. To choose a subset of variables from a larger set according to which original variables have the highest correlations with the principal component factors.
3. To generate a set of factors to be treated as uncorrelated items as one approach to handling multicollinearity in such procedures as multiple regression.
4. To validate a scale or index by representing that its constituent items load on the same factor, and to remove proposed scale items which cross-load on more than one factor.

Pallant (2005) explains that exploratory factor analysis comprises three major stages: (a) assessment of suitability of data for factor analysis, (b) factor extraction, and (c) factor rotation. Therefore, prior to performing the factor analysis, some assumptions should be included in the preliminary analysis performed to check for suitability of the data set for conducting EFA. Table 3.5 presents the summary of these assumptions. For checking of linearity (linear relationship of variables), Hair et al. (2010) suggest the use of P-P plots to check the relationship. The plotted points need to be close to the ideal line for linearity to exist. The issue of multicollinearity, which refers to the degree to which a variable's effects can be predicted or accounted for by the other variables in the analysis, is checked using the variance inflating factor (VIF) and tolerance. The VIF (the inverse of tolerance) values greater than 10 indicate multicollinearity.

Table 3.5: Summary of EFA Requirements on Data Set

Condition	Requirement	Reference
Outliers	No Outliers accepted	(Hair et al., 2010)
Linearity	No Multicollinearity; VIF < 10	(Hair et al., 2010)
Normality	Should be Normally Distributed	(Hair et al., 2010)
Sample Size	Minimum: 5 Cases to each study item	(Pallant, 2005; Tabachnick & Fidell, 2007)
Bartlett's Test of Sphericity	Be Significant ($p < .05$)	(Tabachnick & Fidell, 2007)
Kaiser-Meyer-Olkin (KMO) Index	≥ 0.5	(Hair et al., 2010; Malhotra, 2007)

For checking the normality of the data set, the skewness and kurtosis values are used. According to Tabachnick and Fidell (2007), skewness refers to the symmetry of a distribution, that is, a variable whose mean is not in the centre of the distribution is regarded as a skewed variable, whereas kurtosis relates to the peakedness of a distribution. A distribution is said to be normal when the values of skewness and kurtosis are equal to zero (Tabachnick & Fidell, 2007). The recommended range of skewness values is –1 to +1 (Hair et al., 2010) and for Kurtosis the range is -2.0 to +2.0 (Coakes & Steed, 2003). The issue of sample size was discussed in section 3.10.2. The sample of 508 is greater than the sample needed (400) for 80 items.

The Kaiser-Myer-Olkin (KMO) is a measure of sampling adequacy. The KMO, which is recognized as one of the best measures of determining the suitability of a set of data for subsequent factor analysis, was used to investigate the data in order to decide whether a factor analysis should be undertaken. The Kaiser-Meyer-Olkin measure of

sampling adequacy tests whether the partial correlations among variables are small. Measured by the Kaiser-Meyer-Olkin statistics, sampling adequacy predicts if data are likely to factor well, based on correlation and partial correlation. The Kaiser-Myer-Olkin varies from 0 to 1.0 and small values of KMO suggest that a factor analysis should not be undertaken. The Kaiser-Myer-Olkin should be 0.60 or higher to proceed with factor analysis (Tabachnick & Fidell, 2007).

Bartlett's test of sphericity assesses the overall significance of the correlation matrix. The best result in this test is when the value of the test statistics for sphericity is large and the significance level is small (Nunnally, 1978). Tabachnick and Fidell (2007) point out that data is factorable when the Bartlett's test of sphericity is significant ($p < .05$).

The preliminary analysis of EFA leads to factor extraction. The main goal of factor extraction is to determine the factors that can summarize the interrelationships between the variables. Seven methods of factor extraction are available: principal components analysis (PCA), unweighted least squares, generalized least squares, maximum likelihood, principal axis factoring, alpha factoring, and image factoring (Hair et al., 2010; Malhotra, 2007; Tabachnick &Fidell, 2007).

Although there are many types of extraction techniques, the most frequently used is principal components. Tabachnick and Fidell (2007) suggest that if a researcher is interested in an empirical summary rather than a theoretical solution, PCA is a better choice. The PCA method was preferred for this study than the other methods as this study is interested in an empirical summary rather than a theoretical solution. The items with

factor loading above the cut-off point $\left|0.50\right|$ are retained for further analysis (Hair et al., 2010).

For determining the number of factors to be retained in a scale, one can either retain all factors whose eigenvalues exceed a particular value or retain a certain number of factors. Using the Kaiser's criterion, only components with an eigenvalue of more than 1.0 were selected for further investigation (Hair et al., 2010).

Once the numbers of components have been identified, the next step is to determine the pattern of loadings for easy interpretation. Rotation is used in factor analysis to achieve a simpler factor structure. There are two main approaches for rotation: orthogonal and oblique. The varimax rotation (orthogonal) was used in this study because the major objective of varimax rotation is to have a factor structure in which each variable loads highly on one and only one factor. In other words, the goal of this technique is to simplify factors by maximizing the variance of the loadings within factors, across variables (Tabachnick & Fidell, 2007).

3.12.3.2 Techniques used in confirmatory factor analysis

Confirmatory factor analysis (CFA) is a special case of the structural equation model (SEM). Before the assessment of the measurement model, the process of item purification is applied through multiple iterations of CFA, with the maximum likelihood estimation (MLE) method. Unfitted items are deleted from the measurement model. As recommended by Hair et al. (2010), modification of the initially hypothesized model is performed where it is seen to be applicable. This is achieved based on such indicators as modification indices

(MI), standardized residuals, path estimates, squared multiple correlations, and qualitative review. These model diagnostics are used to suggest model changes, which are known by specification search, whereby an empirical trial-and-error approach is used (Hair et al., 2010). The cut-off points in the CFA process are presented in Table 3.6. Table 3.7 shows the types of fit measures and the thresholds recommended by Hair et al. (2010)

Table 3.6: Model Diagnostics in Confirmatory Factor Analysis

Model Diagnostic	Requirement	Reference
Modification Index (MI)	≥4	(Hair et al., 2010)
Standardized Residuals	< | 2.5 | no problem > | 4.0 | possible problem	(Hair et al., 2010)
Path Estimates (Construct to Indicator)	≥0.5; ideally ≥0.7; and be significant	(Hair et al., 2010)
Squared Multiple Correlations (SMC) or Reliability	≥0.3	(Hair et al., 2010)

Table 3.7: Model Fit Indices

Fit Indexes	Acceptable level
Chi-square	Values with non significant *p*-value
Normed Chi-square (CMIN/DF)	< 3.00
Goodness-of-fit Index (GFI)	≥ 0.90
Adjusted Goodness-of-fit Index (AGFI)	≥ 0.90
Tucker-Lewis Index (TLI)	≥ 0.90
Comparative Fit Index (CFI)	≥ 0.90
Root Mean Square of Error of Estimation (RMSEA)	≤ 0.08

3.12.3.3 Techniques for assessing measurement model

Measurement models are commonly used to assess the construct validity (Churchill, 1979). Construct validity involves the evaluation of the degree to which a measure correctly measures what it is supposed to measure (Cavana et al., 2008; Chen & Paulraj, 2004; Hair et al., 2010; Malhotra, 2007). To achieve construct validity, some conditions must be satisfied including: content/face validity, unidimensionality, reliability, convergent validity, and discriminant validity (O'Leary-Kelly & Vokurka, 1998). Content/face validity does not require statistical tests and was discussed in section 3.11.2. Since content/face validity is established, the unidimensionality and reliability are tested followed by the convergent validity and discriminant validity tests.

3.12.3.3.1 Unidimensionality

Unidimensionality means that a set of items can be explained by a single underlying construct (Hair et al., 2010, p. 696). The procedure for assessing unidimensionality requires that the items are significantly associated with an underlying construct, plus each item being associated with one, and only one, latent variable (O'Leary-Kelly &Vokurka, 1998). By using EFA, the indicator variables load on only one construct with a factor loading of ±0.5. By using CFA, the regression weights are 0.5 or higher with their significant *t*-values (*t*-value $\geq$ 1.96 at a=0.05), as recommended by Hair et al. (2010, pp. 117; pp.708).

3.12.3.3.2 Reliability

Reliability refers to the extent to which measures are free from error, thus, being able to create consistency between measurements of a variable (Hair et al., 2010, p.125; O'Leary-Kelly & Vokurka, 1998; Pallant, 2005). To achieve a good reliability, the

reliability coefficient or Cronbach's alpha should be .7 or higher (Hair et al., 2010, p. 125). By using CFA, the construct reliability (CR) is used, which is also known as composite reliability. Composite reliability refers to the internal consistency of indicators measuring the underlying factors (Fornell & Larcker, 1981). The rule of thumb of CR is that 0.7 or higher implies good reliability (Hair et al., 2010, p. 710). The construct reliability is given in smartPLS software output and can also be calculated using the following equation:

$$CR = \frac{(\sum_{i=1}^{n} \mathrm{Li})^2}{(\sum_{i=1}^{n} \mathrm{Li})^2 + (\sum_{i=1}^{n} \mathrm{Ei})}$$

Where:

L = the standardized factor loading,

i = the number of the corresponding item,

E = the error variance term for an item.

3.12.3.3.3 Convergent validity

Convergent validity refers to the extent to which instruments designed to measure the same construct are related to each other (Malhotra, 2007, p. 279).To examine the convergent validity, the average variance extracted (AVE) was computed by the indicators corresponding to each of the study constructs. The average variance extracted is also given in smartPLS software output and can be calculated using the following equation:

$$\mathrm{AVE} = \frac{\sum_{i=1}^{n} \mathrm{Li}^2}{n}$$

Where:

L = the standardized factor loading,

i = the number of the item.

AVE is the amount of variance that is captured by the construct in relation to the amount of variance due to measurement error. If AVE is less than 0.50, the variance due to measurement error is larger than the variance captured by the construct, and the validity of the individual indicators, as well as the construct, is questionable (Fornell & Larcker, 1981). Thus, convergent validity is established if the AVE for each construct accounts for 0.50 or more of the total variance as applied by Battor and Battour[2] (2010c). Also, this study examined convergent validity by examining the loading paths of all items, which should be statistically significant and exceed 0.50 (Hair et al., 2010, p. 709). It should be noted that reliability is also considered an indicator of good convergent validity if CR is 0.7 or more.

3.12.3.3.4 Discriminant validity

Discriminant validity is the extent to which the measure is indeed novel and not simply a reflection of some other variable (Churchill, 1979). It is the extent to which measures of the constructs are distinctly different from each other. Discriminant validity was examined using the procedure recommended by Fornell and Larcker (1981), whereby discriminant validity is established for a construct if its AVE is larger than its shared variance with any other construct. The AVE was compared with the highest variance that each construct shares with the other constructs in the model.

3.12.3.4 Techniques for structural model assessment and hypothesis testing

The partial least square (PLS) technique is selected to assess the structural model in the current study. Partial least square (PLS), such as smartPLS software, and structural

[2]One of the researcher's papers published in the Journal of Marketing Management.

equation modelling (SEM), such as AMOS software, are second generation data analysis techniques (Gefen et al., 2000). Although there are some diversities between the PLS and SEM programs, the basic specification of the structural model is similar (Hair et al., 2010, p. 775).

SEM is a covariance-based approach using model fitting to compare the researcher's model, as given by theory, to a best possible model fit. In other words, the indices and residuals provided tell how closely the proposed model fits the data as opposed to a best fitting covariance structure (Gefen et al., 2000). SEM reduces the difference between the covariance and those predicated by the theoretical model using a maximum likelihood (ML) function (Chin & Newsted, 1999). Thus, SEM is more focused on explanation and is a more appropriate tool for theory testing (Hair et al., 2010, p. 776).

In contrast, PLS is designed to explain variance, variance-based, similar to OLS multiple regressions (Gefen et al., 2000). Therefore, the focus is much more on prediction (Hair et al., 2010, p. 776). PLS estimates the parameter such that it minimizes the residual variance of all the dependent variables in the model, rather than estimating the variance of all observed variables as in covariance-based SEM(Gefen et al., 2000). Although PLS can be used for confirming the theory, it can also be used to suggest whether relationships exist or not and to propose suggestions for further testing (Chin, 1998). In general, PLS is a prediction model (Chin, 1998; Chin & Newsted, 1999). Table 3.8 shows the comparative analysis between PLS and SEM conducted by Gefen et al. (2000).

It should be noted that PLS, as a method of analysis, is not employed enough in tourism motive studies. The applications of this technique in this study will give some

guidelines to its use in this context. The use of PLS methodology has become an increasingly popular technique in empirical research in international marketing, which may signify an appreciation of the unique methodological features of PLS (Henseler et al., 2009). Henseler et al. (2009) reported that "*As of March 2008, more than 30 articles on international marketing using PLS were published in double-blind reviewed journals*". Moreover, PLS may be considered the method of choice for successful factor studies in marketing (Albers, 2009;cited in Henseler et al., 2009) and for estimating the various customer satisfaction index models (Fornell, 1992).

Table 3.8: Comparative Analysis between PLS and SEM

Issue	**PLS**	**SEM**
Objective of overall analysis	Reject a set of path-specific null hypotheses of no effect.	Show that the null hypothesis of the entire proposed model is plausible, while rejecting path-specific null hypotheses of no effect
Objective of variance analysis	Variance explanation (high R square)	Overall model fit
Theory base	Does not necessarily require a sound theory base. Supports both exploratory and confirmatory research	Requires a sound theory base. Supports confirmatory research
Assumed distribution	Relatively robust to deviations from multivariate distribution	Multivariate normal, if estimation is through ML. Deviations from multivariate normal are supported with other estimation techniques
Sample size	At least 10 times the number of items in the most complex constructs.	At least 100-150 cases

Adopted from Gefen et al. (2000)

PLS has been increasingly used as an alternative to SEM (Hair et al., 2010, p. 775). Thus, the major reasons for using PLS in the current study comparing with SEM are as follows: (a) PLS is more suitable for exploratory studies such as the current study, where some measures are new and the relationships have not been previously tested enough (Ainuddin et al., 2007; Hair et al., 2010; Holzmüller & Kasper, 1991; Lee et al., 2006; Tsang, 2002); (b) PLS is recommended for complex models focusing on prediction and latent variable modelling of interaction effects (Chin et al., 2003), therefore, it is necessary to use PLS for the proposed model as it includes a moderating variable; (c) PLS is recommended when multi-item measures are not available for latent constructs (Hair et al., 2010, p. 778), similar to the current study, because the overall tourist satisfaction variable only has four items and the destination loyalty variable only has four items; (d) PLS is suggested where relationships might or might not exist (Chin, 1998) and where theory is insufficiently grounded (Acedo & Jones, 2007); (e) At an early stage of model development, the regression based approach of PLS is considered more suitable than covariance-based methods such as SEM (Venaik et al., 2005); (f) A model with no strong relationships between variables, such as the proposed model in the current study, is better to be analysed using PLS (Falk & Miller, 1992); and (g) PLS determines the relationships between established indicators to its respective latent variables, which is critical for validating exploratory models (Julien & Ramangalahy, 2003; Mahmood et al., 2004).

The main difference with this discrepancy in objectives between SEM and PLS is that no proper overall goodness-of-fit measures exist for models estimated using PLS. As reported by Tenenhaus et al. (2005, p. 173), *"... differently from SEM-ML, PLS path modeling does not optimize any scalar function so that it naturally lacks of an index that can provide the user with a global validation of the model (as it is instead with $\chi 2$ and*

related measures in SEM-ML). The GoF represents an operational solution to this problem as it may be meant as an index for validating the PLS model globally". Evaluating goodness-of-fit (GoF) can be achieved by calculating the geometric mean of the average communality and the average R^2as using the following equation:

$$GoF = \sqrt{[(\text{average communality}) \times (\text{average } R^2)]}$$

Where:

Average communality is a weighted average of the different communalities with the weights being the number of indicators per latent variable.

In order to examine the relationships between constructs in the proposed model as well as to test the hypotheses, two stages of analysis were performed to evaluate the structural model: (1) Structural model without moderating variable, and (2) Structural model with moderating variable. The method of examining competing models was applied in these two stages as recommended by previous researchers (Chin et al., 2003; Sarkar et al., 2001; Walter et al., 2006). Path modelling and analysis was performed using SmartPLS software (http://www.smartpls.com/).

The path coefficients and the R^2 values derived from the competing models provide the statistical basis for hypotheses testing to determine whether the hypothesized relationships are statistically significant. Path coefficients reflect the strength of the relationships between the exogenous and endogenous variables. The R^2 value indicates the predictive power of a model for the endogenous variables. The significance of the paths is determined by calculating the *t*-value using a bootstrap resampling method (500

samples).The Bootstrap is a general statistical technique for assessing uncertainty through re-sampling data with data replacement (Chin, 1998).

Communality and redundancy coefficients can also be used in essentially the same way as the R^2, since they reflect the relative amount of the explained variance for latent and manifest variables. The blindfolding approach proposed by Wold (1982) was applied to calculate the CV-communality (CV refer to construct cross-validated) and CV-redundancy indexes. The CV-communality index (H^2) measures the quality of the measurement model, whereas the CV-redundancy index (F^2) measures the quality of the structural model (Tenenhaus et al., 2005, p. 174).

According to Tenenhaus et al. (2005, p. 174), the CV-communality index (H^2) has been known as a cross-validated R^2 between the block manifest variables and their own latent variable. The CV-communality measures the capability of the path model to predict the manifest variables directly from their own latent variable by cross-validation. By using the mean of the CV-communality indexes, the measurement model quality can be established if they are positive for all blocks of variables.

According to Tenenhaus et al. (2005, p. 181), the quality of each structural equation is given by the CV-redundancy index defined as a kind of cross-validated R^2 between the manifest variables of an endogenous latent variable and all the manifest variables associated with the latent variables explaining the endogenous latent variable, using the estimated structural model. In other words, it measures the capability of the path model to predict the endogenous manifest variables indirectly from a prediction of their own latent

variable using the related structural relation by cross-validation. By using the mean of the various CV-redundancy indexes (F^2), the structural model quality can be established if they are positive for all endogenous blocks.

The effect size (f^2) is also calculated, which is a measure of the strength of the theoretical relationship including the moderating effect (Chin et al., 2003). The f^2 values of .02, .15, and .35 are considered small, moderate and large effects, respectively (Cohen, 1988);f^2 is calculated by the following equation:

$$f^2 = \frac{R^2 \text{ (interaction model)} - R^2 \text{ (main effects model)}}{1 - R^2 \text{ (main effects model)}}$$

The interaction effect is a result of multiplying the independent and the moderator variable, and to decrease the possible problems with multicollinearity resulting from interaction terms, the mean was centred to the indicators prior to multiplying them, as suggested by Chin et al. (2003). The significance of the effect size is tested by the F statistic, as suggested by Tabachnick and Fidell (2007, p.148). After calculating the F statistic, it is compared with tabled *F*, with k and (N-k-1) degrees of freedom (Tabachnick & Fidell, 2007, p.944). If the obtained *F*> tabled *F*, that means the effect size is statistically significant. The F statistic is calculated using the following equation:

$$F = \frac{R^2/k}{1 - R^2/(N - k - 1)}$$

Where:

K: is the total number of independent variables

N: sample size

3.13 Chapter Summary

This chapter presents a detailed discussion of the chosen research design, methodology and the data collection and analysis methods used to conduct the research study. These issues are addressed in light of the research objectives identified in Chapter One, as well as the literature review in Chapter Two. This research used the two-phase sequential method to achieve the study objectives. Phase one used qualitative approaches – FGDs and individual interviews –in data collection; the second phase used a quantitative approach, namely, the survey. The process of developing the IAD construct and the main statistical analysis techniques are also discussed. Exploratory Factor Analysis and CFA and partial least square (PLS) have been used to test the research framework and the proposed relationships. The following chapters provide the study's results, discuss implications of the findings and suggest directions for further research.