CHAPTER IV

RESEARCH METHODOLOGY

4.1. Overview

This chapter specifies the procedures on how the model has been structured in subsequent chapter. The methodology part includes: research design, population and sample and data collection method (including instruments and procedures). In addition data analysis and ethical considerations are also discussed.

4.2. Research Process

The research process summarises the various steps taken in this study as shown in Figure 4-1. The first step represents the area of specialisation of the research. This is followed by extensive literature review from previous chapters on the research area. The discussion of the theoretical and conceptual framework was established. The next step was pilot study studies, preliminary data testing and adaptation of the instruments. The third steps was field work activities. The final steps will be data analysis, interpretation, discussions and conclusions. These steps were used in attaining the following specific research objectives:

1. To determine whether entrepreneurship training moderates opportunity seeking and readiness towards new venture creation;

2. To determine whether entrepreneurship training moderates motivation and readiness towards new venture creation;
3. To determine whether entrepreneurship training moderates resources and readiness towards new venture creation;

4. To determine whether entrepreneurship training moderates ability and readiness towards new venture creation.

Figure 4-1: Research Process

Step 1
- Introduction: Area
- Identification of research issue
- Literature Review
- Theoretical/conceptual framework
- Hypotheses and Instruments Development

Step 2
- MODEL I: Direct Effect
  To investigate the determinant factors that can influence new venture creation using objective measure (Questionnaire) through SEM.
- MODEL II: Moderating Effect
  To investigate whether entrepreneurship trainings do moderate the relationship of the factors on new venture creation using objective measure (Questionnaire) through SEM.

Step 3
- PILOT STUDY
- FIELD WORK (OBTAINING REAL DATA)
- DATA ANALYSIS

Step 4
- DISCUSSION OF THE RESULTS
- CONCLUSIONS
4.3. Research Design

The approach of the study is hypothetical deductive. The study settings was non-contrived settings since type of investigation was correlations study. The unit of analysis was at individual level because the study involved testing the effect of the entrepreneurial training on students. The time horizon was cross sectional because the effect was tested at a point in time. The sampling frame was drawn from the six faculties in the main campus of USIM.

The survey design was quantitative and deductive in approach. The sampling design was probability sampling design. Simple Random sampling was used because the training involved different faculties in the University. Survey was carried out as the primary data collection method. In order to conduct a survey, there must be a questionnaire established properly. For the purpose of this study, the questionnaire generally started with a brief introduction and then followed by instruction to the respondents in answering the questionnaire.

4.4. Population and Sample

The respondents of this research paper were undergraduates of Universiti Sains Islam Malaysia (USIM). USIM was used in order to encourage more venture creation among Muslim students in Malaysia. The population of USIM has over 10,000 Muslim students. The principles of USIM on building both the revealed and human knowledge makes the University to serve as a perfect sample target of the study. The target respondents was unrestricted to undergraduate's students in USIM. The study operates at 5% confidence interval and 95% confidence level to measure the accuracy and reliability of the data. The minimum sample size for population of 10,000 respondent is 370 (Sekaran, 2003). Therefore, the sample size for this study comprised
of 490 respondents regardless of the age, faculty and course that gave sufficient information for later data analysis purpose. Out of which 418 participated in entrepreneurship programmes. 490 Students provided more accurate information for this research study since it involved training at a point in time.

The sampling technique was drawn in the form of simple random sampling. The random sampling was the forms of probability sampling design where everyone have a known and an equal chance to be selected. Each student was given questionnaire irrespective of faculties until the 590 questionnaires were fully distributed. This sampling design has the least bias and offers the most generalizability. The population of USIM was quite large in number, and it was difficult and due to time constraints to identify everyone in the population.

4.5. Data Collection: Instruments and Procedures

Data for this research study was collected in the form of primary and secondary data. For the primary data, structured questions was adopted. Questionnaires were distributed to the target respondents which were randomly chosen within USIM campus. The questionnaire (see Appendix C) was used to obtain the needed data for analysis. The questionnaire was adopted from Keat, et al. (2011); Choo & Wong (2006); Ummi (2010) and Jumaat, et al. (2013). The questionnaire consisted of two parts. The first part had consisted of questions on demographic profile of the respondents (1-10). The second part consisted of questions eliciting information about readiness, training, opportunity seeking, motivation, resources, and ability. The part II aspect was divided into Section A-F as shown thus.
Table 4-1: Adopted Questionnaire from Previous Studies

<table>
<thead>
<tr>
<th>SECTION</th>
<th>Construct</th>
<th>Questionnaire number</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>section (A)</td>
<td>Readiness towards New Venture Creation</td>
<td>11-17</td>
<td>Keat, et. al. (2011)</td>
</tr>
<tr>
<td>section (B)</td>
<td>Entrepreneurship Training</td>
<td>18-21</td>
<td>Keat, et. al. (2011)</td>
</tr>
<tr>
<td>section (C)</td>
<td>Opportunity Seeking</td>
<td>22-32</td>
<td>Keat, et. al. (2011)</td>
</tr>
<tr>
<td>section (D)</td>
<td>Motivation</td>
<td>33-41</td>
<td>Choo &amp; Wong (2006)</td>
</tr>
<tr>
<td>Section (E)</td>
<td>Resources</td>
<td>42-48</td>
<td>Ummi (2010)</td>
</tr>
<tr>
<td>Section (F)</td>
<td>Ability</td>
<td>49-55</td>
<td>Jumaat, et. al. (2013)</td>
</tr>
</tbody>
</table>

The section (A) part (1-5) was based on Readiness towards New Venture Creation (11-17). The section (B) part (18-21) was based on Entrepreneurship Training/Education. The section (C) part (22-32) was based on Opportunity Seeking. The section (D) part (33-41) was on motivation of students. The section (E) part (42-48) was based on Resources. The section (F) part (49-55) was based on Entrepreneurial Ability of the students. Other aspects of the questionnaire were on feedback from entrepreneurship training programme and recommendation for future entrepreneurship course. This questionnaire was used to achieve the interest of this study.

As for the mode of data collection, self-administered surveys that allowed the respondents to complete the survey was adopted. This was attractive in terms of costs saving and gave respondents control pace at which the questions was filled and avoided interviewer evaluation apprehension (Sekaran, 2003:229-230).

The measurement and measures that was used were based on nominal and interval scale. The rating scale that was used in this research was likert scale (5-point scale). The respondents were asked to state their agreement and disagreement on a 5-point likert scale with 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree and 5 = strongly agree. The likert scale (5-point scale) questions was set based on Readiness towards
new venture creation, entrepreneurship training, opportunity identification, motivation, resources and ability.

This study discussed the past research as the foundation of the research. The secondary data was gathered from Internet, books, magazines, newspapers and also journals. Among all, journal played the most important role as it contributed much on completing the foundation of this study and also literature review, which was presented in previous chapter. Relevant journals were fully utilized to obtain the necessary data for the study.

4.5.1. Pretesting for the Model

In this study, the pretesting was carried out through peer review, expert review and pilot study before proceeding to the main analysis.

4.5.1.1. Peer Review

In the beginning, a draft of questionnaire was tested by three postgraduate students to check the ambiguity in the questionnaire and clarity of the instruments. After receiving their suggestions and corrections the instruments were further reviewed by experts.

4.5.1.2. Expert Review

The experts checked the validity of the questions with suggestions for minor corrections. All the questions on clarity and design of the questions were taken into considerations and further improvements were made. Part of the expert suggested that feedback and recommendation need to be considered in the questionnaire. Also the expert also suggested that demographic characteristics should be included because it
will make the research study to be unique. This was quickly taken care of before the pilot study was carried out.

4.5.1.3. Pilot Test

The pilot study represents the final stage in questionnaire development processes. It is very essential to carry out the pilot test in order to ensure that respondent fully understand the earlier version of the instruments used based on the context of the study (Neuman, 2003). Pilot test for this study was administered through self-administered surveys to 100 undergraduates’ students of USIM from various faculties. Only 50 respondents managed to complete the questionnaires and returned them. There were some confusing questions that were misunderstood by the students and some questions were redrafted to meet the clear understanding of the students. The questionnaire numbering was not consistent. Therefore, the numbering was corrected and the questions were adequately organised. Pilot study of 100 sample size was carried out in order to check the accuracy and reliability of the questionnaire. Some of the questions were redrafted to suit the understanding of students. The questions were increased from 40 to 55 questions through inclusion of demographic characteristics. Also, feedback from the training and recommendation were given by the students. In the process of carrying out the pilot study the following were tested:

a. Descriptive Analysis and Feel for Data

Pilot study was tested as outcomes of the analysis was needed. The data was felt by checking the central tendency and the dispersion through testing using SPSS 21.0. The mean, frequency, range, standard deviation and percentage of frequency and variance in the data gave a good idea of how the respondent had reacted to the items in the questionnaire and how good the items and measures were. The graph, chart and
diagram was widely exercised in disclosing the frequency distribution for information interpretation. Table was used to interpret the other information. Hence, getting a feel for data became the necessary first step in all data analysis.

b. Goodness of Data

Further detailed analysis was done to test the goodness of the data. This was the indication of the stability and consistency with which the instrument measures the concept and helps to assess the "goodness" of a measure. Cronbach's alpha was used as an estimator of the internal consistency reliability of a psychometric test score for a sample of the study. Cronbach’s alpha is a reliability coefficient that indicates how well the items in a set are positively correlated to one another. The closer the Cronbach’s alpha is to 1, the higher the internal consistency reliability. This test was applied on the independent variables with the dependent variable in this study. All the items that could not contribute to the Cronbach’s alpha value were deleted.

c. Factor Analysis

In this study, factorial validity was established through factor analysis. The results of factor analysis was confirm whether or not the theorized dimension emerged. Therefore factor analysis revealed whether the dimensions were indeed tapped by the items in the measure, as theorized. All the items that have low factor loading were deleted before proceeding to the main survey of the study.

4.6. Data Analysis Techniques

After the data have been collected from the respondents, this study used two statistical tools to analyse the data. This was to interpret and analysed the result of the study. The first tool is the Statistical Package for the Social Sciences, (SPSS) version 21.0.
The SPSS 21.0 was used to determine the descriptive analyses such as mean, standard deviations, charts, and other frequencies. The data was also cleaned through SPSS in order to identify and assess the percentage of the missing items. The second statistical techniques used was Maximum Likelihood Estimator (MLE) approach through Structural Equation Modelling using AMOS 21.0 software (James, 2011 and Chumney, 2012). The MLE analysis was used to test the measurement model such as the validity, reliability and structural model used to test the hypotheses of the study. In order to analyse survey data, this study used maximum likelihood Estimators (MLE) approach through structural equation modelling (SEM). SEM overcomes the limitations of Ordinary Least Square (OLS) or First-generation techniques such as Logistic regression, Multiple Regression, Hierarchical regression and others because of its ability to evaluate systems of variables and relationships simultaneously (Chumney, 2012). The SEM method as an extension OLS was developed to cater for the limitation of OLS (Zainudin, 2012). Therefore, SEM is recognized as Second generation statistical technique that is used to analyse interrelationships among multiple variables and simultaneous modelling of relationships among constructs (Gefen, 2000).

4.6.1. Maximum Likelihood Estimator (MLE)

The study chose the MLE analysis to establish the measurement model within the context of the structured models. Due to large samples inherent in this research, SEM technique (i.e., maximum likelihood estimation) is equipped to address this problem. SEM regression models attempt to model the relationships among theoretical constructs and explain the underlying constructs responsible for producing observed
variables. MLE is one of the most common and widely used methods for estimation structural equation models and is available within most SEM software (Kline, 2005).

Maximum Likelihood Estimator (MLE) yields estimates that seek to maximize the likelihood that the observed data come from a population consistent with the implied model. An advantage of MLE is that it is a full-information method, which means that all of a model's parameter estimates are calculated simultaneously (Chumney, 2012). The fit function of an estimation method is the statistical criterion the method aims to minimize; in MLE, the fit function is the difference in covariance between the observed data and the population data specified by the model being estimated.

Two advantages of MLE are that it is scale free (parameter estimates will not change when a variable is transformed linearly) and scale invariant (the fit function is independent of the scale of response data). Despite these advantages, MLE relies on several assumptions. Specifically, MLE relies on asymptotic theory, which implies large samples, and assumes correct model specification, independent observations, independent exogenous variables (i.e., values obtained for exogenous variables are independent), exogenous observed variables measured without error, and the distribution of scores for endogenous variables in the population are multivariate normal (Kline, 2005). Speaking generally, a small sample is problematic in the context of MLE because the estimates and fit tests it produces are asymptotically true (Lee & Song, 2004). This means that without large samples, the validity of statistical inferences may be rightly questioned. MLE is known to be robust to minor violations of its assumptions, but the extent of that robustness varies with the data and model.

MLE is appropriate for this research since the study is based on Confirmatory modelling. PLS is often viewed as more appropriate for exploratory work than for
confirmatory factor analysis (CFA) modeling, and the coefficients it predicts are often consistent but biased compared to other estimation methods such as MLE (Cassell, et al., 1999).

4.6.2. Stages and Processes of Maximum Likelihood Estimator (MLE)

There are two processes involved in the analysis and interpretation of MLE model (Chumney, 2012). First, the measurement model must be validated by evaluating its reliability and validity. The CFA measurement shows the relationship between items and constructs. The second aspect is the measurement of structural equation model where the relationship between the constructs are tested. MLE are explained further thus:

4.6.2.1. Validating the CFA Measurement Model of a Latent Construct

SEM is a confirmatory method providing a comprehensive means for assessing and modifying the measurement model of a latent construct (James, 2011 and Zainudin, 2012). The procedure is called Confirmatory Factor Analysis (CFA). The method has the ability to assess the unidimensionality, validity and reliability of the measurement model (construct).

The study confirmed the unidimensionality, validity, and reliability for all latent constructs involved in the study before modelling their inter-relationship in a structural model (SEM). However, the unidimensionality assessment was examined first prior to assessing validity and reliability.

With Confirmatory Factor Analysis (CFA), the items that do not fit the measurement model due to low factor loading was removed from the model. The study performed the CFA for all latent constructs involved in the model. The study run the CFA for
every measurement model using a pooled CFA. However, the pooled CFA procedure was more preferred for this study than individual. Thus, this procedure for model assessment demonstrated and the assessment for each element was done as follows:

a. **Unidimensionality**

Unidimensionality was achieved when the measuring items had acceptable factor loadings for the respective latent construct. In order to ensure unidimensionality of a measurement model, any item with a low factor loading was deleted. According to Zainudin (2012), for a newly developed items, the factor loading for an item should be 0.5 or higher and for already established items, the factor loading for an item should be 0.6 or higher.

b. **Validity**

Validity is the ability of instrument to measure what it supposed to be measured for a construct. For this study, three types of validity were required for each measurement model:

i) **Convergent validity.** This validity is achieved when all items in a measurement model are statistically significant. The convergent validity could also be verified by computing the Average Variance Extracted (AVE) for every construct. The value of AVE should be 0.5 or higher for the Convergent Validity to achieve.

ii) **Construct validity.** This validity is achieved when the Fitness Indexes for a construct achieved the required level. The fitness indexes and the level of requirement are presented in Table 4-1. In SEM, there are several Fitness Indexes that reflect how fit is the model to the data at hand. However there is no consensus agreement among researchers which fitness indexes to use. Hair, et al. (2010) and Holmes-Smith (2006)
recommend the use of at least one fitness index from each category of model fit. There are three model fit categories namely absolute fit, incremental fit, and parsimonious fit. The choice of index to choose from each category to report depends on which literature is being referred. The information concerning the model fit category, their level of acceptance, and comments are presented in Table 4-2.

<table>
<thead>
<tr>
<th>Name of category</th>
<th>Name of index</th>
<th>Level of acceptance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. absolute fit</td>
<td>Chisq</td>
<td>p&gt;0.05</td>
</tr>
<tr>
<td></td>
<td>RMSEA</td>
<td>RMSEA&lt;0.08</td>
</tr>
<tr>
<td></td>
<td>GFI</td>
<td>GFI&gt;0.90</td>
</tr>
<tr>
<td>2. Incremental fit</td>
<td>AGFI</td>
<td>AGFI&gt;0.90</td>
</tr>
<tr>
<td></td>
<td>CFI</td>
<td>CFI&gt;0.90</td>
</tr>
<tr>
<td></td>
<td>TLI</td>
<td>TLI&gt;0.90</td>
</tr>
<tr>
<td></td>
<td>NFI</td>
<td>NFI&gt;0.90</td>
</tr>
<tr>
<td>3. Parsimonious fit</td>
<td>Chisq/df</td>
<td>Chi square/df&lt; 5.0</td>
</tr>
</tbody>
</table>

*Source: Adopted from Zainudin (2012)*

The acceptable cut-off values reported by researchers may vary depending on literatures that are referred to. One could ignore the absolute fit index of minimum discrepancy chi-square if the sample size obtained for the study is greater than 200 (Hair et al., 2006; Joreskog & Sorbom, 1996). However, the following table presents the literature support for the widely employed fitness indexes.

<table>
<thead>
<tr>
<th>Name of category</th>
<th>Name of index</th>
<th>Index full name</th>
<th>Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. absolute fit</td>
<td>Chisq</td>
<td>Discrepancy chi square</td>
<td>Wheaton et al. (1977)</td>
</tr>
<tr>
<td></td>
<td>RMSEA</td>
<td>Root mean square of error approximation</td>
<td>Browne &amp; Cudeck (1993)</td>
</tr>
<tr>
<td></td>
<td>GFI</td>
<td>Goodness of fit index</td>
<td>Joreskog &amp; Sorbon (1984)</td>
</tr>
<tr>
<td>2. Incremental fit</td>
<td>AGFI</td>
<td>Adjusted goodness of fit index</td>
<td>Tanaka &amp; Huba (1985)</td>
</tr>
<tr>
<td></td>
<td>CFI</td>
<td>Comparative fit index</td>
<td>Bentler (1990)</td>
</tr>
<tr>
<td></td>
<td>TLI</td>
<td>Tucker-Lewis Index</td>
<td>Bentler &amp; Bonett (1980)</td>
</tr>
<tr>
<td></td>
<td>NFI</td>
<td>Normed Fit index</td>
<td>Bollen (1989)</td>
</tr>
<tr>
<td>3. Parsimonious fit</td>
<td>Chisq/df</td>
<td>Chi square/degree of freedom</td>
<td>Marsh &amp; Hocevar (1985)</td>
</tr>
</tbody>
</table>
iii) Discriminant validity. This validity is achieved when the measurement model is free from redundant items. AMOS identifies the pair of redundant items in the model in term of high Modification Indices (MI). This shows some of the weakness of SPSS. The study could delete one of the items and run the model. The study could also set the correlated pair as “free parameter estimate”. Another requirement for discriminant validity is the correlation among exogenous constructs should be less than 0.85.

c. Reliability

Reliability is the extent of how reliable is the said measurement model in measuring the intended latent construct. The assessment for reliability for a measurement model of this study was carried out using the following procedures.

i) Internal reliability: This reliability is achieved when the Cronbach’s Alpha coefficient is greater than 0.7 (calculated in SPSS 21.0).

ii) Composite Reliability (CR): This is the measure of reliability and internal consistency for a latent construct. A value of CR > 0.6 is required in order to achieve composite reliability for a construct. (CR is calculated using the given formula).

\[ CR = \frac{(\sum K)^2}{(\sum K^2) + (\sum 1 - K^2)} \]

where \( k \) = factor loading for every item and \( n \) = number of items in a model.

iii) Average Variance Extracted: This is the average percentage of variation as explained by the measuring items for a construct. An AVE > 0.5 is required (AVE is calculated using the given formula).

\[ AVE = \frac{\sum K^2}{n} \]

where \( k \) = factor loading for every item and \( n \) = number of items in a model.
4.6.2.2. Measurement of Structural Equation Modelling (Multi Group
Regressions Path Analysis)

In this study, the conceptual frame work consists of four constructs in the independent variables. The four constructs of the independent variables are opportunity seeking, Motivation, Resources and Ability. The SEM multi group regression path analysis was applied in interpreting the data collected from the questionnaire which related to the factors that influence the readiness of students towards new venture creation. By this test, the appropriate relationship among these factors was established.

a. The Steps Involved In CFA for the Measurement Model of a Latent Construct

The various steps in CFA for the measurement model of a latent construct (Zainudin, 2012) as used in this study are outlined thus:

- Run Confirmatory Factor Analysis (CFA) for the measurement model
- Examine the required Fitness Indexes for the measurement model Refer to Table 4-1. If the Fitness Index does not meet the required level, examine factor loading
- Delete an item with factor loading less than 0.6 (suggested by the literatures)
- Delete one item at a time (select the lowest factor loading to delete first)
- Run this new measurement model (the model after an item is deleted), examine the Fitness Indexes – repeat step 3-5 until the fitness indexes achieved, if the Fitness Index is still not achieved, look at the Modification Indices (MI), high value of MI (above 15) indicates there are redundant items in the model.
To solve the redundant items, the study choose one of the following:

Choice 1: Delete one of the item (choose the lower factor loading) and run the measurement model and repeat the above steps

Choice 2: Set the pair of redundant item as “free parameter estimate” and run the measurement model and repeat the above steps

- Obtain the Cronbach’s Alpha, Composite Reliability (CR), and Average Variance Extracted (AVE) for the final measurement model
- Report the normality assessment for all measurement models involved

Note: The acceptable value of factor loading, the level of fitness indexes, and also the method of modification to the measurement model varies among the literatures (Zainudin, 2012). All these steps were used to carry out the CFA measure for the latent constructs used in this study. After these steps was carried out successfully the next steps was to turn the conceptual framework to structural equation model and the steps involves are discussed next.

b. Path Analysis Estimates: Hypothesis Testing of Structural Equation Modelling

After validating the CFA measurement model of a latent construct, the unidimensionality, validity and reliability has been achieved. The out of range and missing responses are clean up. The goodness of the measures are established. The hypothesis developed in the study will be tested through at 5% level of significant using SEM through direct and moderating effects.
i. **Analyzing the Direct Effect for Latent Constructs: The Multi-Group CFA**

In this study, the direct effect was performed by using measurement of the pooled constructs. The correlations of the items were determined. The path Arrow which represents the “beta” was shown. The hypothesis of the four constructs were examined. The fitness index was achieved and standardized regression weight as shown by Zainudin (2012). The p-value was used to test the level of significant and the beta sign show the relationships among the constructs (James, 2011 and Zainudin, 2012).

ii. **Analysing the Moderator for Latent Constructs: The Multi-Group CFA**

According to Cohen & Cohen (1983), moderation takes place when the independent variable and the moderating variable have mutual effects on variance of dependent variable than that explained by the direct effect. Entrepreneurship training was used in this study as a moderator because when it is highly effective, it raises students’ inclination and attitude towards entrepreneurship (Souitaris, et al., 2007). The various ways in which moderating effects of latent constructs can be measured are outlined thus as shown from past studies (James, 2011 and Zainudin, 2012). There are few steps involved in performing Multi-Group CFA:

i. Split data into two groups based on the moderator variable to be tested.

ii. Save data into two separate files: Name the files as dataset 1 and dataset 2.

iii. Select the path of interest in the model to test the moderator variable.

iv. Develop two separate AMOS models: Rename as model 1 and model 2.

v. In Model 1, constraint the parameter in the path of interest to be equal to 1.

vi. Name model 1 as the constrained model.

vii. In model 2, do not constrain the relationship in the path of interest.
viii. Name model 2 as the unconstrained model.

ix. Use dataset 1: Estimate the constrained model

x. Use the same dataset 1: Estimate the unconstrained model

xi. Obtain the difference in Chi-Square value between the constrained and the unconstrained model. If the value differs by more than 3.84, then the moderation occurs in that path.

xii. Repeat the same procedure using dataset 2.

xiii. Use dataset 2: Estimate the constrained model

xiv. Use the same dataset 2: Estimate the unconstrained model

xv. Obtain the difference in Chi-Square value between the constrained and the unconstrained model. If the value differs by more than 3.84, then the moderation occurs in that path.

Note: Data set 1=High Trained Group Data; Data set 2= Low Trained Group Data

4.7. Summary

Chapter 4 discussed the research design, population and sample, data collection instruments and data analysis techniques. The research design such as the sampling design, unit of analysis and external interference were examined, sources of data as well as the instrumentation used were presented in this chapter. The data analysis techniques is divided into pretesting and post testing techniques. The results of the statistical analysis and hypothesis testing will be presented in the next chapter.