CHAPTER III
RESEARCH METHODOLOGY

3.1 Introduction

This chapter discusses the methodology of this research. Amongst others, this chapter will elaborate on the study's conceptual model, research design, the population and sample of the research, as well as data collection procedure. This chapter also reports on the pilot test done for the measurement of the students' acceptance model. This chapter ends with a discussion of the statistical techniques that were used to analyze the data.

3.2 Research Model

As the foundation of the problems of the study, the research framework provides illustration that all ideas, meanings and propositions are related to the research problems. This study proposed that albeit the existence of modern technological advancements, the Technology Acceptance Model (TAM) still holds itself as the primary reference tool in numerous bodies of research since it was first established in 1989 (Davis, 1989). In fact, TAM still maintains its reputation as an efficient mechanism to gauge acceptance. Still, the demand for assessment techniques that fit the process of technology – namely m-learning - evaluation cannot be ignored. Additionally, a more comprehensive study is crucial so that the various environmental influences could be addressed (Chen, 2010). Also, even though the usage of TAM has seen progression, there are still issues with regard to conducting a more thorough evaluation in creating the method of applying the model in the framework of m-learning as well as its method of measurement.

Another point worth mentioning is that, although factors such as service quality, social influence, culture, compatibility, perceived usefulness, and perceived ease of use were present in the previous studies, none of them were used in any of the studies' completed model. Thus, the complex relationships among these factors and the behavioral intention with regard to the utilisation of m-learning among higher education institutions have not been evaluated and ascertained. Owing to this, utilising a more complete model, the study will present new ideas for implementing m-learning to students. The aforesaid model is presented in Figure 3.1.

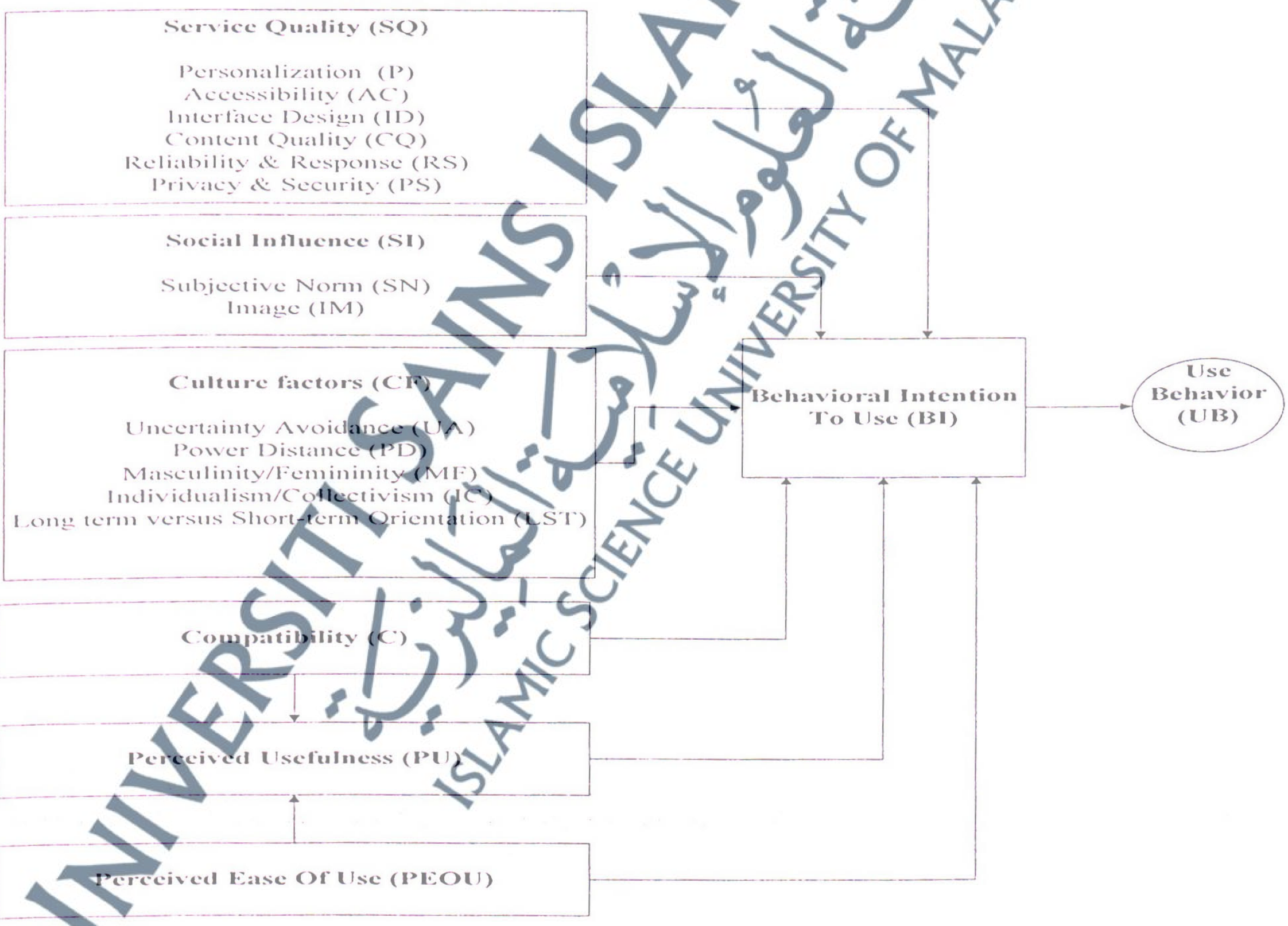

Figure 3.1: Research Model

3.3 Research Instrument

This section entails deliberations on the instrument of research that is adopted for this study. Specifically, this section highlights the development of instrument for gathering data on the students' demographics section. In addition to that, the perceptions of m-learning as well as mobile technologies are also evaluated. Further, the awareness and requirements of m-learning services in the final data collection on the students' acceptance of m-learning service are gauged through the measurement of the factors of service quality, social influence, culture, compatibility, perceived usefulness, perceived ease of use, behavioral Intention to use and use behavior.

3.3.1 Instrument Development

Quality of data is highly influenced by the instrument developed (Eccles et al., 2011), and the instrument used for obtaining data in this study is that is constructed based on the guidelines proposed in (Gay et al., 2006). It is important to note that in order to be effective, questionnaire developed should appear attractive to participants. Aside from that, it should be brief meaning that, it contains only the items that appropriately represent the objectives of the research. Further, the questionnaire should be formulated to collect only the necessary demographic information, focus on items based on particular topics or ideas, define and explain terms that are ambiguous, use clear words and free from leading questions. Also, the items presented should be organized from general to specific, and items should be grouped together with their response options, and lastly, the questionnaire should be pilot tested. In addition to that, researcher should also place careful attention on

the questionnaire`s length, order, content, and also, the type of each question. The questionnaire developed in this study is an adaptation of questionnaires from prior studies. The details are shown in Table 3.1 and Table 3.2.

Table 3.1: Constructs and Variables of Research Model

Construct	Sources
Social Influence (SI)	Davis et al., (1989) & Ajzen, (1991).
Subjective Norm (SN)	
Image (IM)	
Service Quality (SQ)	Al-Mushasha & Hassan, (2009) & Abbitt, (2006).
Personalization (P)	
Accessibility (AC)	
Interface Design (ID)	
Content Quality (CQ)	
Reliability & Response (RS)	
Privacy & Security (PS)	
Culture factors (CF)	Li et al., (2009) & Lee et al., (2008).
Uncertainty Avoidance (UA)	
Power Distance (PD)	
Masculinity/Femininity (MF)	
Individualism/Collectivism (IC)	
Long term versus Short-term Orientation (LST)	
Compatibility (C)	Moore & Benbasat (1991); Chen et al., (2004) Tung et al.,(2014)
Perceived Usefulness (PU)	Davis et al., (1989).
Perceived Ease Of Use (PEOU)	Davis et al., (1989).
Behavioral Intention To Use (BI)	Venkatesh et al., (2003).
Use Behavior (UB)	Chen et al., (2004).

Table 3.2 : Research Students' Awareness and Requirements of Mobile Learning Services

Construct	Sources
Awareness of mobile learning service aspects	Walton et al., (2005) & Karim et al., (2006).
Current access to learning resources	Walton et al., (2005).
Mobile technologies for learning services	Walton et al., (2005) &Kim et al., (2006).
Limitations of mobile technologies	Walton et al., (2005) &Kim et al., (2006).
Suggested university mobile services	Karim et al., (2006).

3.4 Research Hypotheses

This study intends to set up a hypothetical model that could offer explanation and prediction on student acceptance and usage of m-learning services in the setting of higher education. In the formation of the model, the TAM's and IDT's belief - intention – behavior relationship will be used. Specifically, this relationship postulates that students' actual of m-learning use is instantaneously controlled by their behavioral intention to use (BI) (Venkatesh et al., 2003).

This study intends to investigate student acceptance of m-learning services by examining the student's intention to use and the use behaviour of the m-learning services. This notion is chosen because this study is interested in looking at acceptance as well as its level, and to achieve this, both data of use behavior and the intention to use will be

evaluated. Another reason for using these data is that, the data of use behaviour provides good indication of use continuation in the future, and this is important in the context of m-learning. Parthasarathy and Bhattacherjee (1998), indicated that online service utilization means continuance in service adoption and for this reason, this study employs both intention to use and behaviour usage gauge student acceptance of m-learning.

The model of this study extends TAM and IDT's concept of belief through the inclusion of three more constructs namely, social influence, perceived service quality and culture. The addition of perceived service quality signifies the service-oriented aspect of m-learning, while the addition of perception tackles students' general concern with regard to mobile technology and the internet in general. Then, model proposes that the service quality, social influence, culture, compatibility, perceived usefulness and perceived ease of use shape the attitude of a student with regard to m-learning. As such, this study develops the hypotheses below:

H1: A student's behavioral intention to use m-learning service has effect on use behavior of the m-learning services.

H2: A student's service quality of m-learning service has a direct effect on behavioral intention to use the m-learning service.

H3: A social influence of m-learning service has a direct effect on behavioral intention to use the m-learning service.

H4: A student's culture of m-learning service has a direct effect on behavioral intention to use the m-learning service.

H5: The compatibility has a direct effect on perceived usefulness of the m-learning service.

H6. The compatibility has a direct effect on behavioral intention to use the m-learning service.

H7: A student's perceived Ease of Use of m-learning service has a direct effect on behavioral intention to use the m-learning service.

H8: A student's perceived ease of use of m-learning service has a direct effect on perceived usefulness of the m-learning service (PEOU → PU).

H9: A student's perceived usefulness of m-learning service has a direct effect on behavioral intention to use the m-learning service (PU → BI).

All hypotheses are interrelated and they are the building blocks of the model proposed by this study. The model proposed by the study is called Mobile Learning Acceptance Model and it is illustrated in Figure 3.2.

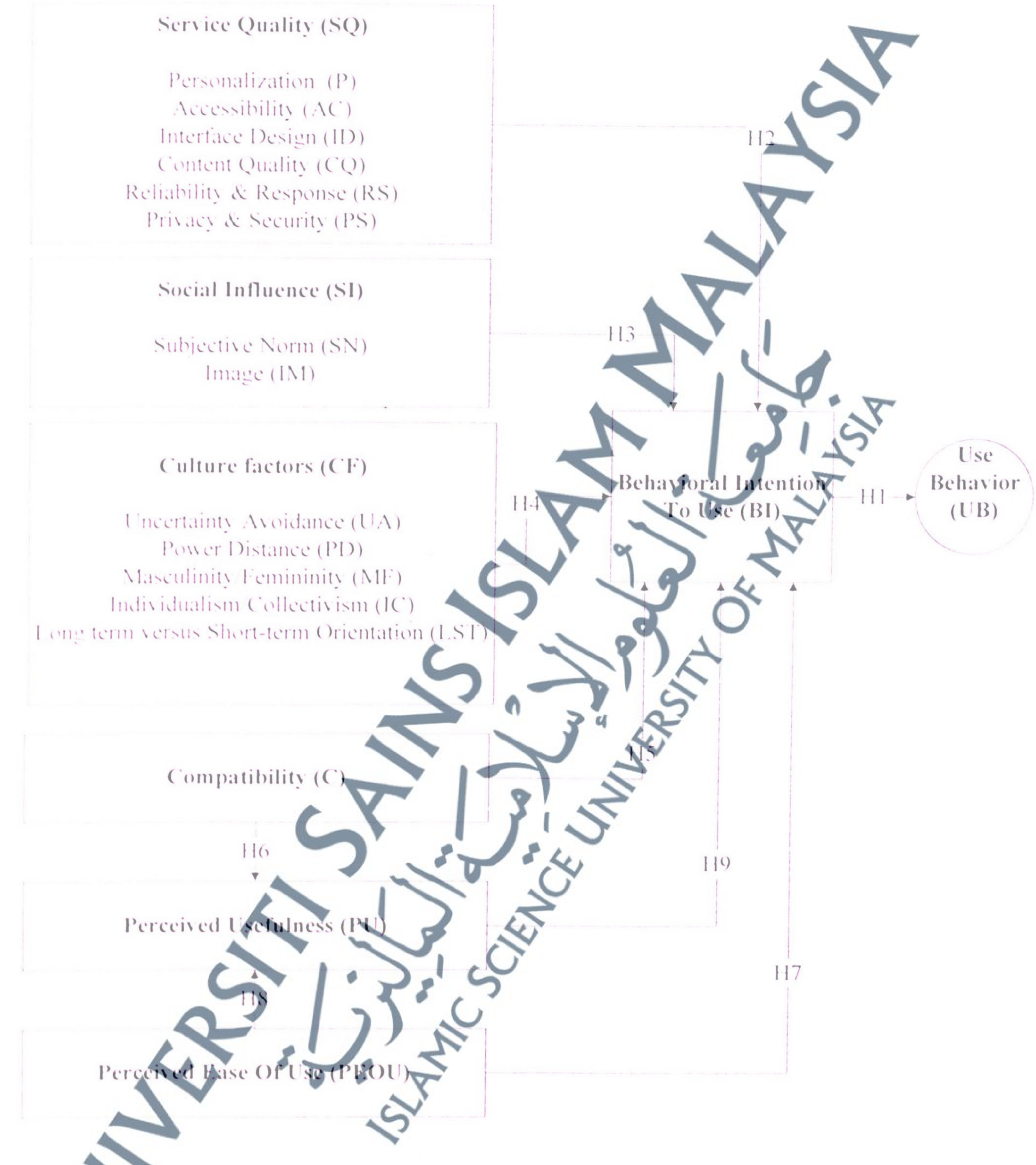

Figure 3.2: Research Model Hypotheses

3.5 Research Design and Approach

Research is referred as a formal, systematic application of the scientific method to the study of problems (Gay et al., 2006). In detail, the scientific method entails formulating hypotheses based on observation, hypotheses implications deductions, implications testing, and hypotheses confirmation or disconfirmation. On the other hand, a hypothesis entails descriptions of particular behaviors, phenomena, or events that happen. This study will apply the inductive reasoning in formulating hypotheses based on the past research and literatures. With regard to the research design, its purpose is to allow the research problem to be examined and this is achieved by formulating a research process that produces specific answer to particular questions. There are generally two main research approaches which are quantitative and qualitative approach (Oppenheim, 1998), and the application of each approach in research depends on the disposition and the prerequisite of the research questions as well as research objectives. Pertaining to the research types, Yin (2003) has classified them according to the purpose of research and the types are exploratory, descriptive or explanatory research. In the context of this study, a quantitative descriptive survey design will be used. Specifically, quantitative research entails the collection and analyses of numerical data in order to explain, predict, and or control a phenomenon of interest (Gay et al., 2006).

This study has two different surveys for different research objectives. Therefore, for each questionnaire, there is a specific data collection procedure that is presented in detail. While data collection procedure for students' awareness and requirements of m-learning services is presented in section 3.4.1 and data collection procedure for the model of the

adoption and use of m-learning services is presented in section 3.4.1. It is worth noting that, for research questions or research objectives that revolve around beliefs or behaviors, survey methods would be appropriate (Neuman, 2003) and this view is supported by Zikmund (2003) who establishes surveys as better methods for gauging respondents' awareness, opinions, and behavior. Then, as indicated by Davis et al., (1989), ascertaining the students' acceptance towards a particular system will entail measuring their opinions and behavior, and therefore the primary strategy of this research is to adopt a survey design in order to gain the necessary information from the selected sample.

3.6 Population and Sampling Method

Castillo (2009), indicated that research population generally entails large collection of individuals or objects being the primary focus of a scientific query. Additionally, it is important to note that research is conducted to benefit the population itself. Testing every individual in the population could generate accurate results however, the size of the population could be too large that testing each individual would be costly and time consuming. Therefore, as indicated by Sekaran (2003), relying on sample techniques would be more practical. Here, the size of the sample would be determined and it depends on the factors such as the degree of accuracy required, the number of variables present in the study and the statistical tools used.

3.6.1 Sampling Frame

This study has selected all undergraduate and postgraduate students of Islamic universities in Malaysia enrolled in the 2014/2015 academic year as the target population. Specially, two universities are selected which are the International Islamic University of Malaysia (IIUM), and Universiti Sains Islam Malaysia (USIM). The details are presented in Table 3.3.

Table 3.3: Distribution of the two Islamic Universities in Malaysia

University	City
International Islamic University Malaysia (IIUM)	Kuala Lumpur
Universiti Sains Islam Malaysia (USIM)	Nilai

As members of the population are homogenous, this study applied the cluster with Probabilities Proportional to Size (PPS). Two public universities have been selected in the study and the universities are the clusters. The selection of universities is the first step of sample selection and the number of undergraduate and postgraduate students from each selected university is shown in Table 3.4.

Table 3.4: Number of Students from each University

University	Population of Students
International Islamic University Malaysia (IIUM)	32.086
Universiti Sains Islam Malaysia (USIM)	13.022
Total	45.108

Sources: (Malaysia Educational Statistics, 2014)

Secondly, proportionate random sampling was used to determine the number of students that formed the sample scope for the current study (Table 3.5). The number of students from the International Islamic University Malaysia (IIUM) is the largest number with about 71 % of the total number of undergraduate and postgraduate students that formed the sample scope for the current study. Also, Universiti Sains Islam Malaysia (USIM) has 29%.

Table 3.5: Proportion of Universities Students Sample with the Corresponding Percentage

University	Population of Students	% of Sampling
International Islamic University Malaysia (IIUM)	32,086	71%
Universiti Sains Islam Malaysia (USIM)	13,022	29%
Total	45,108	100%

3.6.2 Sample Size

Sampling entails a process of deciding on the participants for the purpose of a study by way that they exemplify the whole population that they belong to (Gay et al., 2006). As indicated by Bless and Higson (1995), determining a sufficient sample size of which will represent appropriate quantity of the whole population is crucial, and a population of more than or equal to 40000 would require a sample of at least 380 (Krejcie and Morgan, 1970, as cited in Sekaran, 2003). Thus, based on the aforementioned, it is appropriate to select a sample of at least 381 students based on the size of the target population of the

research. In addition, this study employs the proportionate random sampling in order to determine the number of students who made up the study's sample scope, as shown in Table 3.6.

Table 3.6: The Sample Distribution on each University based on its Percentage from entire Population

University	Percentage from Target Population	Population of Students	Allocated Sample	Systematic Random Sampling
International Islamic University Malaysia (IIUM)	71%	32,086	271	341
Universiti Sains Islam Malaysia (USIM)	29%	13,022	110	153
Total	100%	45,108	381	

3.6.3 Systematic Sampling Design

According to Sekaran (2003) the systematic sampling design involves drawing every n^{th} element in the population starting with randomly chosen element between 1 and n. In this study the researcher chose a random sample by which 381 respondents were systematically identified from the two universities in Malaysia as represented in Table 3.6. The list of students from each university was used to ensure randomness.

3.7 Questionnaire Design

The questionnaire used in this study is comprised of 98 items divided into three major sections and the original language of the questionnaire is English. The questionnaire is attached in Appendix A.

3.7.1 Initial Instrument Structure

Drafting instructions that are clear and understandable is the very critical step prior to organizing the survey questions. In order to achieve this, a cover page comprising the title and a brief description of the research is also included in the questionnaire. In addition to that, respondents are also supplied with a consent page that provides information with regard to anonymity and confidentiality, procedure and risks, respondent's rights, the expected time needed to complete the questionnaire, the description of offered incentive and also contact information. On the whole, the instrument consists of nine-sections as described below:

Section 1 includes general information of the respondents namely, gender, age, education background, study program, mobile device ownership, type of mobile devices owned, provider of mobile service, experience in mobile applications and type of wireless connections used.

Section 2 encompasses dimension A, through dimension E.

Dimension A consists of items on the awareness of the names of mobile technologies. A list containing six items is provided for the respondents to identify their awareness of mobile technology. The scale used is a five point Likert scale ranging from 1-5, where 1= Not Aware, 2= Somewhat Aware, 3= Not Sure, 4= Aware and 5= Very Aware. For each item, respondents are required to select the scale that is most appropriate to them, that is, with regard to the degree to which mobile technology as perceived by them, is important or not important for them as students.

Dimension B consists of six items that measure the available learning recourses such as online learning care. The respondents are asked to indicate their view on the access of online learning resources using the five point Likert scale ranging from 1-5, where 1= Not Aware, 2= Somewhat Aware, 3= Not Sure, 4= Aware and 5= Very Aware. For each item, respondents are required to select the scale that is most appropriate to them.

Dimension C measures the respondents' views on the use of mobile technologies for learning services. This dimension is represented by six items and the scale used is the five point Likert scale ranging from 1-5, where 1= Strongly Disagree, 2= Disagree, 3= Neutral, 4= Agree and 5= Strongly Agree.

Dimension D comprises of thirteen items that gauge on the respondents' views on the limitation of mobile technologies. The items use the five point Likert scale ranging from 1-5, where 1= Strongly Disagree, 2= Disagree, 3= Neutral, 4= Agree and 5= Strongly Agree, and respondents are required to select the scale where appropriate.

Dimension E comprises of items that are formulated to determine the m-learning services that are often used in the higher education setting. These items make up of eight services

that respondents could have access to, at their universities. In addition to that, aside from the eight items listed, respondents could also add in other mobile services that they may use. In this section, a five point Likert scale is used ranging from 1-5, Where 1= Never 2= Seldom 3= Sometimes, 4= Often and 5= Always. In answering this section, respondents are required to indicate the degree to which services were important or not important for them as students.

Section 3 comprises of Subsection 1 through Subsection 8.

Subsection 1 consists of items that were formulated to measure the service quality (SQ) of m-learning services in the higher education setting. There are twenty one items in this subsection and they were specifically formulated to gauge the performance of using m-learning services in the higher education environment. Six dimensions are reflected in this subsection of service quality evaluation namely, accessibility, interface design, reliability & response, content quality personalization, and privacy & security. In terms the scale used, it is the five point Likert scale ranging from 1-5 where 1= Strongly Disagree, 2= Disagree, 3= Neutral, 4= Agree and 5= Strongly Agree.

Subsection 2 consists of five items covering the social influence (SI) to use m-learning services in the higher education setting. The purpose of these items is to gauge the performance of using m-learning services in the environment of higher education. Two dimensions are reflected in this subsection namely, social influence and subjective norm image. With regard to the scale, the five point Likert scale is used ranging from 1-5, where 1= Strongly Disagree, 2= Disagree, 3= Neutral, 4= Agree and 5= Strongly Agree.

Subsection 3 consists of items covering the culture factors (CF) of m-learning services in the higher education setting. There are fifteen items that were formulated to gauge the performance of using m-learning services in the higher education environment. Subsection reflects five dimensions on cultural differences namely, Uncertainty Avoidance, Power Distance, Masculinity/Femininity, Individualism/Collectivism, and Long term versus Short-term Orientation. A scale of Likert's five points is used to measure the respondent's degree of agreeability where 1= Strongly Disagree, 2= Disagree, 3= Neutral, 4= Agree and 5= Strongly Agree.

Subsection 4 consists of three items, which were formulated to measure the compatibility of m-learning services in the setting of higher education. Specifically, these items are used to indicate the degree of compatibility of m-learning services use with most aspects of respondents' education purposes and information searching namely, their lifestyles and their engagement with the higher education environment. To gauge the respondent's degree of agreeability, this subsection utilizes the five point Likert scale ranging from 1-5 where 1= Strongly Disagree, 2= Disagree, 3= Neutral, 4= Agree and 5= Strongly Agree.

Subsection 5 consists of five items that measure the perceived usefulness of using m-learning services in the higher education environment. Specifically, these items were formulated to ascertain the perception of respondents towards usefulness of the use of m-learning services in their higher education environment. The five point Likert scale ranging from 1-5 is used to ascertain the respondent's perception where 1= Strongly Disagree, 2= Disagree, 3= Neutral, 4= Agree and 5= Strongly Agree.

Subsection 6 consists of five items that were formulated to measure the perceived ease of use of m-learning services in the context of higher education. These five items specifically measure the perception of respondents in their use of m-learning services in their higher education environment with regard to the service being 'easy to use.' The five point Likert scale ranging from 1-5 is used where 1= Strongly Disagree, 2= Disagree, 3= Neutral, 4= Agree and 5= Strongly Agree.

Subsection 7 consists of four items that were formulated to measure behavioral intention to use m-learning services in the higher education environment. Specifically, these four items were constructed to gauge the respondents' behavioral intention towards the use of m-learning services in their higher education setting. In this subsection, the five point Likert scale ranging from 1-5 is used where 1= Strongly Disagree, 2= Disagree, 3= Neutral, 4= Agree and 5= Strongly Agree.

Subsection 8 comprises of two items that were formulated to measure use behavior of m-learning services in the higher education environment. The first item is a question on whether the respondents use m-learning services frequently. The scale used for this question is the five point Likert scale ranging from 1-5 where 1=Strongly Disagree, 2=Disagree, 3=Neutral, 4=Agree and 5=Strongly Agree. Meanwhile, the second item is a question that was constructed to measure the frequency of use of m-learning services, specifically, the respondent is asked to indicate the approximate number of times that he/she used m-learning services. In responding to this question, the respondent is presented with four frequency options namely, daily, weekly, monthly and a few times a semester. Both questions in this subsection can be used as alternative measures for usage.

However, using frequency will create a different perspective of usage from the actual number of times of use (Igbaria et al., 1997). As such, in gauging the actual usage, this study decides to use both types of questions.

3.8 Pilot Test

Pilot test is an experimental study conducted to improve certain research instrumentations (Zikmund, 2003) and it is regarded as an important step in developing the measurement scales. Not only that, through a pilot test, researcher could detect weaknesses and the potential failure particularly the instrumentation. As such, this study could benefit from a pilot test in a sense that the pilot test would increase the measurement's accuracy and consistency. In fact, pilot test examines the instrument's content validity and initial reliability; each of these is highlighted in the following sections.

3.8.1 Content Validity

Content validity is the extent to which a test measures the content area that it intends to measure (Gay et al., 2006). Apart from that, content validity is established by expert judgement. The two requirements of content validity are the item validity and sampling validity. Item validity is about whether the test items are relevant in the measurement of the intended content area, while sampling validity is about how well the test represents the overall content area tested. In this study, the researcher presented the questionnaire to three academic experts and three students from Universiti Sains Islam Malaysia (USIM) of reviewers to validate the questionnaire of reviewer. They were asked to give feedback

on the following issues: the length of time needed to complete the instrument, the clarity of the statements and instructions, typographical errors and the layout/structure of the instrument. The suggestions from this pre-testing, three useful feedbacks were obtained. Firstly, a minor change was made to the questionnaire wording of the cover letter. However, no comments were suggested regarding the comprehensibility of the statements and instructions. Secondly, all respondents took approximately 15 minutes to complete questionnaire. The thorough review that was done to the questionnaire was in order to ensure that the questionnaire is adequate, comprehensive, appropriate, comfortable, and has quality and clarity. Additionally, from the reviewers' suggestions, the researcher was able to improve on the questionnaire's arrangement, flow and sequencing.

3.8.2 Instrument Reliability

As mentioned earlier, the pilot test is performed to ascertain the reliability of the measurement instruments prior to the main empirical study. Meanwhile, as documented by Gay and Airasian (2000), reliability entails the extent to which a test consistently measures whatsoever that is being measured. Further, Cronbach's alpha measures the initial reliability of internal consistency of data that were obtained from pilot study (Cronbach, 1984) and the Cronbach alpha value can be increased in either the average correlation or the number of items (Zander & Kogut, 1995). Further, it was indicated by Henryson (1971) that an "item-to-total-test correlation should fall between the range of 0.3 to 0.7 for inclusion" in a survey test. Afterwards, a reliability analysis in a more detailed manner is performed on the full set of data. Then, if reliability coefficient

obtained has high value, it means that the instrument is highly reliable. As a starting point, Devellis (1991) and Nunally (1978) indicated that the minimum satisfactory reliability coefficients vary from 0.70 to 0.80 and more specifically, Nunnally (1978) proposes that the value must be greater than 0.7. In terms of sample size for the pilot test, it should not be less than 30 (Hunt et al., 1982) and as such, the pilot test for this study was conducted on 31 randomly selected students from the Universiti Sains Islam Malaysia (USIM). Table 3.7 and Table 3.8 present the test results. Finally, this study uses the prerequisites presented in this paragraph as a guideline for the reliability test.

Table 3.7: Scale Reliability Alpha – Pilot Test of Model's Questionnaire (n=31)

Variable	N. of Items	Alpha (a)
Personalization(P)	5	.872
Accessibility (AC)	4	.820
Interface Design (ID)	4	.720
Content Quality (CQ)	3	.710
Reliability & Response (RS)	2	.730
Privacy & Security (PS)	2	.721
Subjective Norm (SN)	3	.792
Image (IM)	2	.731
Uncertainty Avoidance (UA)	3	.877
Long term versus Short-term Orientation (LST)	3	.722
Power Distance (PD)	3	.740
Individualism/Collectivism (IC)	3	.710
Masculinity/Femininity (MF)	3	.720
Compatibility (C)	3	.702
Perceived usefulness	5	.790
Perceived Ease of Use (PEoU)	5	.810
Behavioral Intention to Use (BI)	4	.800
Use behavior (UB)	2	.730

Table 3.8: Scale Reliability Alpha – Pilot Test of Students' Awareness of M-Learning (n=31)

Variable	N. of Items	Alpha (a)
Awareness of mobile learning service aspects.	6	.733
Current access to learning resources.	6	.720
Mobile technologies for learning services.	6	.713
Limitations of mobile technologies.	13	.710
Suggested university mobile services.	8	.720

3.9 Final Instrument

The behavior data is collected using the final instrument and this final instrument is formulated based on the results of the validity and initial reliability tests. It should be noted that in developing and testing the instrument the matter of concern was on the instrument with an appropriate design so that methods of statistical analysis can be imposed on, and owing to this, the final instrument retained the sections that were included in the initial instrument. The finalized instrument is attached in Appendix A.

3.10 Data Collection Procedure

This study obtains data from the students of the identified universities through the use of developed questionnaire, and prior to distributing the questionnaire permission was obtained from the Dean of Research and Postgraduate Studies Center in each university. As indicated by Sekaran (2003), if the researcher "knows exactly what is required and how to measure the variables of interest," then, a structured questionnaire is a suitable data collection method. Sekaran (2003) further states two conditions that would make the use of questionnaire an efficient method for data collection and these conditions are: (1) when it is necessary to protect the privacy of the respondents, and (2) when the responses are expected to be standardized. Aside from that, questionnaire is relatively cheaper to administer, familiar to most people, reduces bias, and generally easy to analyze. With regard to the duration of data collection, the two participating universities were given one month to distribute the questionnaires.

3.11 Technique of Data Analysis

This study employs a variety of analysis techniques. Specifically, one way Analysis of Variance (ANOVA) and T-test were used on the demographics data to ascertain if there are any mean differences. Meanwhile, to ascertain the status of direct relationship (whether it is positive or negative), analysis of correlation analysis with Pearson correlation matrix was conducted. The same method is also applied in ascertaining the relationships strength between the posited factors or variables.

Then, to measure the independents variables' (IVs) impacts toward the dependent variables (DVs), the multiple regression analysis was performed. Then, stepwise regression was conducted to find the most substantial predictors' variables imparting influence on m-learning acceptance.

3.11.1 Descriptive Statistics Analysis

Researcher' employ the descriptive statistics to explore, describe and summarize the collected data. Additionally, descriptive statistics is used to portray the various characteristics of data, verify any violation of the principal assumptions for the statistical methods to be used in the study, and address specific research questions (Pallant, 2007). With regard to this study, the measurements are on the statistics of central tendency as well as variation that include range, means, as well as standard deviation.

3.11.2 Factor Analysis

As indicated by Hair et al., (1998), understanding the variables' dimension in the model being proposed or relationships in empirical research is one crucial step in data analysis. Meanwhile, the factor analysis is performed to ascertain the interrelationship (correlation) structure among multitude of items, and as proposed in (Hair et al., 1998), this is accomplished by outlining the common underlying dimensions known as factors. In the context of this study, factor analysis was performed primarily to 1) investigate the construct validity of the measuring concept (Cooper et al., 2003), 2) decrease the amount of variable, and 3) ascertain the structure in the relationship between variables by terming a group of featured dimensions. As such, Hair et al., (1998) suggest that factor analysis can be used to reduce data or detect structure.

Two common types of factor analysis are the Exploratory Factor Analysis (EFA) and the Confirmatory Factor Analysis (CFA). Specifically, the EFA is normally used to summarize a set of structure variables (Coakes et al., 2006) and ascertain the underpinning dimensions of a group of constructs, which are evaluated by a specialized instrumentation. On the other hand, the CFA is applied to affirm a theory pertaining to the structure of a specific domain (Hair et al., 2006; Coaks et al., 2006). As such, the factor analysis techniques conducted in this study should identify and examine the core dimensions of a set of variables, and for this purpose, the EFA is regarded as appropriate.

The EFA applied in this study utilized the principal components method to evaluate the dimensions from the items of the instrument. In particular, the principal components

analysis is a method of factor extraction that is performed to produce the uncorrelated linear combinations of the observed variables. The principal components analysis is variance-based with maximum variance at the first component and as indicated by Coakes and Steed (2003) and Hair et al., (1998), the items removal procedure and factor extraction was iterative for achieving simple factor structure using the four criteria highlighted in the following paragraphs.

First of all, the suitability of continuance using factor analysis is decided by the measure of the Kaiser-Meyer-Olkin (KMO) on adequacy of sampling. The use of the KMO is to ascertain if the size of partial correlations among variables is small; if they are small, then they are prone to factor well. As indicated by George and Mallery (2007), with regard to factor analysis, the value of KMO should be more than 0.60 and thus, if the value yielded is 0.90 or higher then, it can be considered as excellent.

Conversely, the Bartlett's test of sphericity (BTS) entails a test performed to measure multivariate normality. Using BTS, researcher could ascertain whether the correlation matrix is an identity matrix, and if so the factor model can be considered as inappropriate. With regard to this study however, the value for the KMO is higher than 0.6 while the BTS shows significance. Therefore, the utilization of factor analysis in this study was appropriate.

Subsequently, for items that do not load with all other items, they would be omitted. This is done by using the factor matrix of loadings, or correlation analysis can also be performed between the items and factors. In particular, loadings values that are below 0.40 are considered as low, while those values that are higher than 0.6 are regarded as

high (Hair et al., 1998) and therefore, if items have loadings of less than 0.40 they should be removed.

Thirdly, items that load on more than one factor should be eliminated because they violate the simple structure factor solutions, which stipulate only one loading on any factor for each variable because when there is double loading, interpreting the output will be difficult. Double loading occurs when the factor scoring on more than one factor at 0.50 or lesser. In addition to that, as proposed in (Hair et al., 1992), items loaded on a factor should be eliminated if it seems theoretically unwarranted for the particular item to be to be linked with the rest of the items in the factor.

Finally, it is recommended that researcher omit the items that have value of measure of sampling adequacy (MSA) that is lower 0.50 in the anti-image matrix which comprises of the partial correlation coefficients' negative values. An ideal factor model has small off-diagonal elements while the anti-image correlation matrix's diagonal shows measure of sampling adequacy for a variable, where the acceptable level is more than 0.5.

3.11.3 Reliability Analysis

Reliability is, "the extent to which measurements are free from random-error variance" (Hayes, 1998). In other words, if there is random error, reliability of the measurement will decrease. On the other hand, reliability can also be regarded as the degree to which a test consistently measures whatever it is measuring. With regard to a questionnaire, it should have high reliability so that the scores reliably indicate the underpinning dimension.

Hayes (1998) highlighted two advantages of having a highly reliable scale which are, 1) it differentiates between different satisfaction levels in a better way (unlike the low reliability scale) and 2) it is more potent in discovering significant relationships between variables which are truly associated with one another. Simply put, by using a scale that is highly reliable, researcher would be able to identify real dissimilarities regardless of size. On the other hand, researcher would only be able to notice only in large real dissimilarities if they use the low reliability scale. The three common formats of reliability available for research perusal are the internal consistency, test-retest reliability and also equivalent form reliability, and each form may be useful in different types of study.

For a cross-sectional study like the current study, where it is only conducted once, internal consistency is the most suitable; the internal consistency reliability must be tested. Internal consistency reliability is the extent to which items in a single test are consistent among themselves and with the test as a whole, and it can be measured using three different approaches namely, Split-half reliability, Kuder-Richardson, and Cronbach's alpha (Cronbach, 1984). Additionally, as indicated in (Gay et al., 2006), reliability is shown mathematically using a reliability coefficient that is acquired through the use of correlation in which a high reliability is signified by high reliability coefficient. Further, if items in the instrument employ the multiple score choices, for instance, the Likert scales, then the Cronbach's alpha can be used to determine the internal consistency reliability, where, if the items are standardized, then the Cronbach's alpha is based on the average correlation of items within a test, but if the items are not standardized, the Cronbach's alpha is based on the average co-variance among the items. Additionally,

since Cronbach`s alpha can be interpreted as a correlation coefficient, its value ranges from 0 to 1. The reliability of scales is affected by two common factors, which are the amount of items in the scale and the sample in which the reliability is computed. Having more items in the scale means that the instrument is more reliable providing that, the additional items reflect the same concept being measured. Not only that, with regard to the concept being measured, Hayes (1998) stressed that the sample should be heterogeneous. Aside from being influenced by some factors, reliability itself influences the degree of the correlation between any two scales. For instance, scales that have low reliability result in low correlation value and as a result, as indicated by Hayes (1998), it is also probable that in the study, the relationship between two variables is inaccurately concluded. Reliability values obtained in basic research should not be less than 0.80 (the value can definitely be higher) (Nunnally, 1978) but it should be noted that increasing reliabilities exceeding the value of 0.80 will not impact the correlation between scales in a drastic manner. It is widely suggested that the alpha value should not be lower than 0.70 (Netemeyer et al., 2003) but some scholars such as Clark and Watson (1995) contend that higher alpha values of at least 0.80 are required. The criteria for alpha levels by (Devellis 1991) are shown in the Table 3.9.

Table 3.9: The Criteria for Alpha Levels

Value	Ruling
Less than 0.60	Not acceptable
Between 0.60 and 0.65	Not desirable
Between 0.65 and 0.70	Minimally acceptable
Between 0.70 and 0.80	Respectable
Between 0.80 and 0.90	Very good

Source: (Devellis, 1991)

For the purpose of this study, the threshold selected as an acceptable level of reliability is at 0.70 and above.

3.11.4 Correlation Analysis

Pallant (2001) describes correlation analysis as a statistical method used to describe the strength and direction of the linear relationship between two variables. The commonly known is the Pearson's correlation coefficient (r), and it is applied to illustrate the strength and direction of the relationship between two variables.

Correlation is the most common measure of linear relationship and it is suitable for the interval and ratio scales. On the other hand, the coefficient encompasses values ranging from −1 to +1, where the value denotes the strength of the relationship, while the sign (- or +) denotes positive or negative relationship. Compeau and Higgins (1995) proposed the value of 0.05 as the lower limit of substantive regression coefficients. Nonetheless, with regard to substantive correlations, other researchers are tend to prefer the critical value of 0.10 and higher ($r \geq 0.10$). As indicated, the value of the correlation coefficient could range between −1.0 and +1.0, however, it is crucial for researchers to ascertain whether any relationship indicated between two variables is significant (i.e. if the relationship exists merely by chance or it is highly likely that the relationship behaviorally exists). It should be noted that within the researches in social science, the typically accepted significance level is at 0.05. Further, a guideline for describing the strength of relationship between two variables (r) was formulated in (Cohen, 1988) as presented in Table 3.10.

Table 3.10: Cohen's Guideline of Correlation Strength

r values			Strength of relationship
r = +.10 to .29	or	r = -0.10 to -0.29	Small
r = +.30 to .49	or	r = -0.30 to -0.49	Medium
r = +.50 to 1.0	or	r = -0.50 to -0.1.0	Large

Source :(Cohen, 1988)

3.11.5 Regression Analysis

Multiple linear regressions and stepwise regressions were applied in this study. Specifically, as described by Aksornsingchai and Srinilta (2011), multiple regression is a statistical method conducted to model a linear relationship between a dependent variable (predictand) and one or more independent variables (predictors). In other words, it tries to discover the best-fit plane and the fit can be measured by the coefficient of multiple determinations (R^2).

However, prior to conducting the multiple regression analysis, several assumptions must be fulfilled and these assumptions are normality, linearity, multicollinearity and multivariate outliers (Hair et al., 1998). In statistics, as indicated by (Draper and Smith, 1981), stepwise regression entails regression models in which the choices of predictive variables are provided by an automatic procedure. Not only that, stepwise regressions entails the step-by-step iterative formation of a regression model involving automatic selection of independent variables. There are three ways to perform stepwise regression.

The first way is by testing one independent variable at a time and incorporating it in the regression model if it is statistically significant, while the second way is by adding in all the potential independent variables into the model, while omitting variables that are

not significant statistically. Finally, the third way is by combining the first those and second methods. The test statistic (P-value) ranges from 0 to 1 for both multiple regression test and stepwise regressions test, and in the context of this study, it is considered as significant if P value is 0.05 or less (Stengel et al., 2009).

3.11.6 Differences Analysis

T-test is a measure used to ascertain if the mean scores between the two sets of variables indicates difference that is statistically significant with regard to their level of acceptance pertaining to the implementation of m-learning. In this study, the first assumption that was investigated is the homogeneity of variance, and this is done by applying the Levene's test to ascertain variance's equality. Then, if there is violation on the assumptions of equal variances, the t-value reported for equal variances not assumed is chosen. Then, in order to ascertain if any differences exist in the level of acceptance of m-learning among demographic variables with more than two categories, one way analysis of variance (ANOVA) is applied. It should be noted that, since ANOVA test operates on the premise that all variances are equal, researcher performed the Levene's test for homogeneity of variance prior to performing the ANOVA tests to certify that there is no violation on the assumptions of homogeneity of variance.

In addition to that, one way ANOVA is an appropriate tool to test the same hypothesis when comparison is made between two or more groups. Specifically, the hypotheses for the mean comparison are:

H_0: The population means for all the groups are the same.

H_A: The population means for at least two groups are different.

Table 3.11 illustrates the primary data analysis methods that are employed in this study.

Table 3.11: The Data Analysis Techniques Used in the Research

Research Questions	Analysis Techniques
1. What are the user requirements towards the use of m-learning services in the higher education environment?	Descriptive Analysis
2. What are the factors and variables that could influence the acceptance of m-learning services among students?	Pearson product-moment Correlation Multiple regression analysis & Stepwise regression analysis
3. What is the relationship between compatibility and perceived ease of use and perceived usefulness?	Pearson product-moment Correlation & Multiple regression analysis

3.12 Summary

Research design and methodology are highlighted in this chapter. In addition to that, the chapter discusses the subject of pilot study as a procedure to refine the research instrument based on reliability analysis results, together with content validity of the construct. Finally, the chapter outlined the description of data collection procedure as well as the data analysis techniques applied in this study.