DEVELOPMENT OF SITUATIONAL JUDGMENT TEST (SJT) INSTRUMENT TO MEASURE THE EMPLOYABILITY SKILLS OF GRADUATES

MOHD HAFIZ BIN ZAKARIA

MASTER OF SCIENCE (STATISTICS)
UNIVERSITI UTARA MALAYSIA
2013

Permission to Use

In presenting this thesis in fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the University Library may make it freely available for inspection. I further agree that permission for the copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisors or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to:

Dean of Awang Had Salleh Graduate School of Arts and Sciences
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok

Abstrak

Globalisasi membuatkan pengrekrutan graduan untuk sesebuah organisasi menjadi lebih rumit. Majikan percaya bahawa tenaga kerja yang berdaya saing adalah kunci kepada kejayaan sesebuah organisasi. Dalam situasi semasa, walaupun graduan dikatakan mempunyai kemahiran khusus yang baik, mereka masih mempunyai masalah dengan kemahiran kebolehpasaran, dan hal ini menjejaskan kualiti kerja mereka. Tahap kemahiran kebolehpasaran yang baik adalah penting bagi memastikan bahawa graduan bersedia sebelum memasuki dunia pekerjaan. Oleh yang demikian, tahap kemahiran kebolehpasaran mereka perlu dinilai terlebih dahulu. Di dalam kajian ini, instrumen kemahiran kebolehpasaran untuk graduan Universiti Utara Malaysia merentasi bidang pengajian telah dibangunkan menggunakan kaedah penilaian ujian pertimbangan situasi (SJT). Dua belas (12) item telah dibangunkan bagi mengukur kemahiran komunikasi, etika dan moral profesional, kemahiran keusahawanan, pemikiran kritikal dalam penyelesaian masalah dan kualiti kendiri. Kesahihan instrumen telah dicapai melalui pendapat pakar dan kebolehpercayaan (dari segi kestabilan) adalah berdasarkan ujian Chi-Square untuk kehomogenan dan ujian Mann-Whitney U. Seterusnya instrumen yang telah disahkan tersebut digunakan untuk mendapatkan skor norma bagi setiap kemahiran yang diuji menggunakan kaedah *Bias-Corrected and Accelerated* (BCA) *bootstrap*. Skor norma tersebut menunjukkan bahawa responden-responden memiliki tahap etika dan moral profesional yang tinggi; dan tahap yang sederhana bagi kemahiran-kemahiran yang lain. Berdasarkan kepada dapatan kajian, instrumen yang dihasilkan boleh digunakan oleh graduan, universiti, majikan, agensi-agensi kerajaan dan agensi-agensi pembangunan kerjaya apabila menilai tahap kemahiran kebolehpasaran.

Kata kunci: Kemahiran kebolehpasaran, Ujian pertimbangan situasi, *Bias-corrected and accelerated* (BCA) *bootstrap*

Abstract

Globalization makes graduate recruitment for an organization more complex. Employers believe that a competitive workforce is the key to success of an organization. Currently, although graduates are said to have good specific skills, they still have problems with employability skills, and these problems affect the quality of their work. A good level of employability skills is essential in order to ensure that graduates are ready before entering the work market. Therefore, their level of employability skills should be evaluated. In this study, an instrument of employability skills for Universiti Utara Malaysia graduates across various fields was developed by using the assessment approach of situational judgment test (SJT). Twelve (12) items were developed measuring communication skill, professional ethics and morality, entrepreneurial skill, critical thinking in problem solving and personal quality. Instrument's validity was achieved through expert opinion and the reliability (in terms of stability) was based on the Chi-Square for homogeneity test and Mann-Whitney U test. Next, the validated instrument was used to obtain norm scores for each tested skill using the Bias-Corrected and Accelerated (BCA) bootstrap. The norm scores indicated that the respondents possessed a high level of professional ethics and morality; and a moderate level for other tested skills. Based on the findings, the instrument produced can be used by graduates, university, employers, government agencies and career development agencies when evaluating the level of employability skills.

Keyword: Employability skills, Situational judgment test, Bias-corrected and accelerated (BCA) bootstrap

Acknowledgement

Table of Contents

List of Tables

LIST OF FIGURES

List of Appendices

List of Abbreviations

BCA	Bias – Corrected and Accelerated
CA	Communication Apprehension
CGPA	Cumulative Grade Point Average
COSTAM	Confederation of Scientific and Technological Associations in Malaysia
EDA	Exploratory Data Analysis
ICT	Information and Communication Technology
K-S	Kolmogorov-Smirnov
MOHE	Ministry of Higher Education
My3S	Malaysian Soft Skills Scale
NWRC	National Work Readiness Credential
LPQ	Personal Quality Related to Life
CPQ	Personal Quality Related to Work
SJT	Situational Judgment Test
UUM	Universiti Utara Malaysia
WAGE	The Workforce Alliance for Growth in the Economy Certificate Program
WRC	Work Readiness Credential
WSC	Workforce Skills Certification System

CHAPTER ONE
INTRODUCTION

1.1 Research Background

Today, globalization has made the recruitment of workforce in an organization more crucial. An organization that has a competent workforce is said to have a special asset that can help provide a high return on investment and can guarantee the lifetime of the organization (Curtis & McKenzie, 2001; Quek; Armstrong, 2005; Singh & Singh, 2008). In recruiting professionals in the 21st century, employers prefer to hire fresh graduates as they are the group of candidate that is expected to have potentials to help the organization produce fresh ideas and new techniques in enduring changes in globalization (ACNielsen, 2000).

At present, the number of graduates that exceed the number of job vacancies available each year seems to be favourable to employers (Raybould & Sheedy, 2005). This is because they are said to have flexibility in choosing quality graduates that can meet their demand to fill professional positions in their organization. However, despite the abundance of graduates, most of them have low level of work readiness (Quek, 2005). Hence, while employers find it difficult to hire potential employees (Raybould & Sheedy, 2005), many graduates fail to be employed. For those who did get a job, their work performances are not encouraging and have led them to lose their jobs (Zhiwen & Van der Heijden, 2008). A special committee named The Committee to Study, Review and Make Recommendations Concerning the Development and Direction of Higher Education was commissioned by the Ministry of Higher Education Malaysia to study graduate employability. The study indicates that most of the graduates did not meet the requirement of employers.

Among the weaknesses identified are communication, information and knowledge, and self-confidence (Malaysia Ministry of Higher Education [MOHE], 2006a). According to the Department of Statistics (2008), graduate unemployment rate in 2007 in Malaysia was 34.8%. The rate increased to 35.5% in 2010 (Department of Statistics, 2011).

Essentially, to be a competent employee, one needs to possess two types of skills, namely specific skills and employability skills (Singh & Singh, 2008). Specific skills are discipline-specific skills (Bridgstock, 2009; 2011) and is meant to be used in a particular work field. For example, biochemistry graduate should have the ability to apply principles to biochemistry practice in order to conduct laboratory experiments. Employability skills are more generic and can be applied to all types of jobs and positions. For example, common employability skills needed by employers are communication and problem-solving skills (Raybould & Sheedy, 2005). Employability skills are complementary to specific skills. Presently, although graduates are generally good in their specific skills, they seem to have problems in applying them to real situation (Quek, 2005; Singh & Singh 2008). This is because specific skills can only be fully utilized when it is used together with employability skills (Quek, 2005).

Employability skills are the work skills needed in obtaining and maintaining a job (Commonwealth of Australia, 2002). Examples of these skills are the ability of graduates to; 1) Deal with uncertainty, 2) Working under pressure, 3) Plan, 4) Communicate, 5) Deal with information technology, and other skills involving adaptability and intelligence (Raybould & Sheedy, 2005). MOHE (2006b) views

employability skills as fundamental work skills that must be possessed by graduates in global employment.

According to the Malaysian Minister of Higher Education in the 10th Confederation of Scientific and Technological Associations (COSTAM) in July 2000, graduate who lack employability skills may not be able to apply the knowledge gained at university effectively (Quek, 2005). Consistently, graduates who are unable to show a good performance in work are not only incapable in providing any significant contribution; they are even being regarded as a burden and an obstacle to the success of the organization. Hence, in selecting an employee, employability skill is one the important things employers look at (Quek, 2005). Employers generally believe that it is more difficult to train employability skills than specific skills that are easier to develop through training courses (Raybould & Sheedy, 2005).

The above discussion has shown the importance of employability skills to graduates. Without these skills, it would be hard for them to survive in today's working environment. Therefore, their level of employability skills needs to be assessed. Within this assessment, universities will be able to monitor the employability skills of their students prior to their graduation and develop suitable training programs. Not only will the graduates able to secure good employment, they could also help enhance the credibility of their Alma mater among employers. Furthermore, this assessment will benefit the employers in screening candidates. Screening involves two stages ("Career in", 2012). The first stage involves the screening of work application forms. In this stage, applicants will be shortlisted based on pertinent information such as work experience and academic achievement. In the second

screening stage, the shortlisted applicants will be called for an initial interview to obtain the applicants' psychological attribute and skills (Breaugh & Starke, 2000; Guion & Highhouse, 2006; "Career in", 2012).

Assessment of employability skills can be used to assist employers in two ways. Firstly, it can be used after the first but before the second stage. By doing so, the attributes and working skills of the applicants can be assessed earlier and refinement can be made. As a result, the screening process in the second stage will be more effective because only applicants with favourable attributes and acceptable level of working skills can proceed to that stage. The second stage can be treated as the attributes and skills confirmation check point of applicants through a subjective evaluation by employers. This approach can increase the accuracy of selecting a suitable applicant to be hired. Secondly, the assessment can be used as a substitution for the interview in the second screening stage. This is because the objective of the interview and the assessment is the same, that is, to extract information about the attitudes and the working skills possessed by the applicants. In this stage, the applicants will be asked to go through the assessment. With this approach, a more objective judgment can be made and the time and cost associated with the exercise can be reduced.

The availability of graduate employability skills assessment can also benefit government agencies involved in developing training programs for graduates. Typically, skills enhancement modules used in such programs are developed to train and enhance the skills of graduates, regardless of the current level of skills. But by using the assessment, the participants can be clustered based on their level of skills

and the training program can be designed accordingly. This will help reduce the cost incurred on graduate training because graduates will only be trained on the skills they lack.

Finally, the assessment may be beneficial to the career development agencies in the work placement process. Generally, clients (i.e. employers) have already identified what they are looking for in potential employees. Part of the agencies' function is to find and suggest the right candidate to their client. The instrument able to help in identifying the criteria of graduates in their database so that a more accurate suggestion on the suitable candidates can be made. Besides, the result of the assessment may also be compiled in the curriculum vitae of graduates and become a useful reference to the employers in seeking the right employees to be hired.

1.2 Employability Skills Instruments at Present

A widely used mechanism to assess employability skills is work readiness instrument (Alicea & Scott, 2007). Examples of such instrument are the National Work Readiness Credential (NWRC), WorkKeys Career Readiness Credential (WorkKeys), Workforce Skills Certification System (WSCS), the Workforce Alliance for Growth in the Economy Certificate Program (WAGE) (Alicea & Scott, 2007), and Malaysian Soft Skills Scale (My3S) ("Malaysian Soft Skills Scale", 2011).

The National Work Readiness Credential (NWRC) is an instrument to assess those who want to enter bottom-level jobs (non-supervisory, non-managerial, nonprofessional positions) ("FAQs", 2011). The instrument consists of four tests;

oral language test, reading test, math test, and situational judgement test (SJT). The SJT assesses the ability to; cooperate, dealing and negotiating with conflict, observe critically, solve problem and making decision, and be responsible. The first three tests are cognitive-oriented assessment while SJT tests applicants on a hypothetical scenario (Alicea & Scott, 2007).

The WorkKeys Career Readiness Credential (WorkKeys) is an instrument targets on high school students and adults seeking employment or currently employed in entering bottom level (non-supervisory, non-managerial, nonprofessional positions) or middle-level (supervisory position) jobs. The skills tested are reading for information, applied mathematics, locating information, applied technology, writing, business writing, teamwork, observation and listening (Alicea & Scott, 2007). It is an instrument that was developed using the mixed methods of cognitive-oriented and self-reported assessment (ACT, 2008; Alicea & Scott, 2007).

The Workforce Skills Certification System (WSCS) was developed on youth and adults that are entering the workforce, transitioning to work, incumbent workers, and dislocated workers. The skills tested are reading, mathematics, problem solving, critical thinking, teamwork skills, customer service, project development, and presentation. Each skill is tested using cognitive-oriented assessment (Alicea & Scott, 2007; Eguez & Dayton, 2006; "Workforce Skills Certification", 2011).

The Workforce Alliance for Growth in the Economy Certificate Program (WAGE) is an employability skills instrument targeted to those who seek for bottom-level jobs. One-hundred and twelve items are used to test four areas of skills i.e. mathematics,

communication, reading and writing. Beside these tests, applicants are also allowed to take five other optional tests. They are career scope assessment, wage post test, spatial relations test, dexterity test, and computer literacy test (Alicea & Scott; "Adult Education", 2011). In this instrument, cognitive-oriented assessment is used (Alicea & Scott, 2007).

The Malaysian Soft Skills Scale (My3S) is an instrument targeted at students of higher education institutions of Malaysia. The objective of the instrument is to obtain information about the level of employability skills possessed by students. It consists of two-hundred and twenty items that measure communication skill, critical thinking and problem-solving skill, teamwork, continuous learning and information management, entrepreneurial skill, professional ethics and moral, leadership skill, and self-appreciation (MOHE, 2006b; "Malaysian Soft Skills Scale", 2011). Students are to self-perceived their skill level.

In summary, there are two different populations whose are the main targets of the above instruments. For NWRC, WSCS and WAGE, their targeted population is the communities of the United States of America such as youth (high school student) and adults that seek for non-professional jobs (entry-level). As for WorkKeys, its target population is similar to NWRC, WSCS and WAGE, which involve youth and adults. However, instead of conducting an assessment for entry-level jobs, the assessment of WorkKeys is also available for supervisory positions. Finally, My3S was developed specially to measure the employability skills of higher education students in Malaysia.

With regard to employability skills, the four different instruments above are each consist of different number of skills. Seven skills are assessed in WSCS, eight skills in NWRC and My3S, and nine in WorkKeys and WAGE. Finally, with regard to measurement method, three categories of method can be identified. These are self-reported assessment (WorkKeys and My3S), cognitive-oriented assessment (NWRC, WorkKeys, WSCS, and WAGE) and situational judgment test (SJT) (NWRC). Self-reported assessment is a measurement method that relies on respondents' perception and report is commonly made on attitudes (Holbrook, 2008; Salkind, 2006). The cognitive-oriented measurement method is an assessment based on cognitive observation and evaluation of the respondents (Rupp, 2007). Finally, SJT is an assessment based on real situations where the respondents choose the best option to deal with the situations (Joiner, 2008).

In sum, different instruments are; designed for different targeted populations, using various number of employability skills, and use a variety of measurement methods. These three issues are thoroughly discussed in the next section to highlight the gaps between the existing instruments.

1.3 Problem Statement

The problems with the present instrument mentioned above consist of a sheer variety on three issues, which are target population, employability skills, and measurement method.

In this study, the target population comprises graduates of higher education institutions in Malaysia. Because NWRC, WSCS and WAGE focus on youth with

high school education level and adults who seek for bottom-level jobs, these are not suitable for graduates. The same argument may be applied to WorkKeys because it is meant for middle-level jobs only. As for My3S, the respondents are existing university students in Malaysia, not graduates. With regard to the employability skills that need to be measured, the most important skills for graduates are selected. The selection of employability skills for this study is explained in great detail in chapter two (Section 2.3).

Self-reported assessment (used by WorkKeys and My3S) and cognitive-oriented assessment (used by NWRC, WorkKeys, WSCS, and WAGE) proved to be ineffective measurement methods (Lievens et al., 2008). Self-reported assessment assumes that respondents are able to willingly answer the questions honestly and/or accurately. However, in reality, this may not be the case. According to theory of reward and punishment (Denning et al., 1989; Holbrook, 2008), sensitive questions that can affect one's credibility will tend to make the individual build a favourable image of himself or herself to get rewarded. Moreover, from the perspective of psychology, self-evaluation toward one's attitude is based on his/her memory of events that have occurred in the past, and this could lead to less accurate answers (Dunning et al., 1989; Fielding, 2008). Because of this bias tendency , self-reported assessment is less reliable.

Cognitive-oriented assessment, on the other hand, is capable of addressing the limitation of self-reported measurement because the responses are not based on individual perception. Even though it could be valid in measuring abilities, its result may show a sign of biasness when the mean differences of scores in the subgroups of

the measured population were diverse with respect to their demography. This situation is known as adverse impact. Small differences in mean are viewed as having a low adverse impact and large mean differences indicate a high adverse impact (Lievens et al., 2008). Based on a study by Jensen (1998) and Nguyen et al. (2007), cognitive ability test showed a high adverse impact in the mean scores between races of the measured population. This arises because cognitive ability test has high cognitive loading and this causes high adverse impact. Therefore, the result of the cognitive-oriented measurement method could not be standardized across respondents' demography (Lievens et al., 2008). The problems of self-reported and cognitive-oriented assessment are explained in depth in chapter two.

Lievens et al. (2008) state that a good measurement method needs to be efficient, realistic and standardized. By efficient, it means that the measurement is accurate and has a low possibility of response bias. By realistic, the presented question in the instrument portrays a realistic circumstance on the field of interest and, finally, by standardized, the scores obtained by respondents must be systematically consistent despite demographic differences (e.g. age, race, etc.).

SJT (as used by NWRC) is efficient because the response is not based on the perception of respondents while self-reported assessment is (Peeters & Lievens, 2005; Lievens et al., 2008). Secondly, the criterion of realistic measurement is certainly met because SJT is by nature realistic in content. For instance, the question itself is based on a real scenario (Jesiek & Woo, 2011; Lievens et al., 2008). Finally, the adverse impact showed in SJT is low to moderate (Jensen, 1998; Lievens et al., 2008; Whetzel et al., 2007).

The level of cognitive loading in SJT may vary depending on the type of response instruction and the presentation format of the instrument. This is because different response instructions contain different levels of cognitive loading and the same reason can also be applied to the presentation format. For example, Nguyen et al. (2007) found that behavioural tendency response instruction in a video-based SJT is less cognitively loaded than knowledge based response instruction in a written-based SJT. Therefore, the adverse impact in SJT is reduced when behavioural tendency response instruction was used instead of knowledge based response instruction. The same situation also occurred when video SJT was used instead of the written-based SJT. However, in spite of the variation, the instrument of SJT still shows lower adverse impact than cognitive ability test (Lievens et al., 2008; Nguyen et al., 2007) (a detailed explanation on both response instruction and presentation format are offered in chapter three). This means that an instrument developed using SJT is a standardized instrument capable of application across population, hence addressing the gap of a cognitive-oriented assessment.

Based on the criteria of efficiency (low response bias), realism (real issues) and standardization (low adverse impact), SJT is a better measurement to use. Therefore, it is used in developing the instrument to assess graduate employability skills in this study. Besides developing an employability skills instrument, this study also concerned with its generic applicability across graduates. The developed instrument was evaluated on respondents by taking into consideration their fields of study (as implemented by ACNielsen, 2000). In this study, qualitative and quantitative fields of study were evaluated. If the evaluation shows no significant differences exist in

the scores of the respondents, then the instrument can be said to be suitable for all graduates.

Additionally, this study seeks to find a way to extract more meaningful information from the scores obtained by respondents. The scores are intended to rate the level of employability skills possessed. Independently, the score provides little information on the level of employability skills possessed by each respondent. In order to obtain more meaningful information, the norm-reference method was conducted to obtain norm scores. By using the scores, the second objective of this study, which was to obtain the norm score distribution, can be achieved. Next, based on the distribution, individual score can be compared with the average population's score and a more significant information can be obtained (Cohen & Swerdlik, 2009; Colton & Covert, 2007; Oppenheim, 2001). To conclude, in addition to providing information on the individual level of employability skills, the scores can be used to determine the respondents' ranking of employability skills within its population. With this information, the level of employability skills possessed by the respondents will be more comprehensible.

1.4 Research Objectives

The objectives of this study are:

1. To develop an instrument to measure graduate employability skills of graduates.

2. To test the instrument across various fields of study.

3. To obtain a norm score for the target population.

1.5 Significance of the Study

The development of this instrument will help fill the gap of other employability skills instruments by providing a better alternative method to the self-reported and cognitive-oriented assessment. Generally, the researcher expect that the usage of this instrument can benefit five main parties namely universities, government agencies, employers, career development agencies, and graduates. The outcome of the instrument is expected to assist universities in monitoring the level of employability skills possessed by their students. Hence, appropriate skills development programs can be designed. The usage of the instrument can benefit employers at the initial stage of the recruitment process, as explained earlier. If the instrument is used, the screening stage where candidates are shortlisted and the evaluation of the personalities and capabilities of the candidates can be integrated at the screening stage. Therefore, the recruitment process can be optimized.

The instrument is also useful for government agencies. The scores obtained from the implementation of the instrument can be used to determine the cluster of skills that need to be developed. By clustering the skills, training programs can be designed accordingly. This will directly reduce the cost and time required in the graduate training programs without affecting the quality of the training. The development of employability profile of graduates using the scores obtained from this instrument will assist career development agencies in appropriate placement of graduates based on the skills they possess. Finally, for the graduates themselves, the usage of this instrument will help raise awareness on their actual level of employability skills possessed. With this awareness, they could make prepare themselves before entering the workforce. Additional, the result obtained from the instrument can also be an

added value in their curriculum vitae and can be treated as a third party assessment of their skills.

1.6 Research Outline

This thesis is divided into six chapters. The first chapter introduces the current pattern of recruitment by employers, the need for competent workers and the type of employees preferred by employers. It then discusses the present situation of graduate employability, and the potential causes of graduate to be unemployed . The researcher then offer a solution by using a new instrument to assess graduates employability, after limitations of the existing instruments have been highlighted. Finally, the objectives and the significance of the research are explained.

Chapter Two has four main sections. The first main section (Section 2.2) discusses in detail the five employability skills instruments identified in Chapter One. The second main section (Section 2.3) discusses and compares the employability skills considered necessary to graduates. Next, the third main section (Section 2.4) discusses three types of measurement method, namely, self-reported assessment, cognitive-oriented assessment, and SJT. Finally, the fourth main section (Section 2.5) explains the steps of instrument development.

Chapter Three is devoted to SJT. The discussion of SJT is made in five parts: definition, history, properties, development steps, and strength. Chapter Four discusses the methodology. Specifically, it explains the methods as used to achieve the research objectives i.e. to develop an instrument that can effectively measure

graduate employability skills by using SJT, to ensure the generic applicability of the instrument across graduates, and to obtain a norm score distribution.

Chapter Five presents the results and discusses them. The results consist of exploratory data analysis (EDA), hypothesis testing on the differences of the scores based on field of study of the respondents, the scoring process, and the development of norm scores. Chapter Six, which is the final chapter, concludes the research work. Specifically, the chapter describes the research implication, its limitation and offers some suggestions for future research.

CHAPTER TWO
EMPLOYABILITY SKILLS INSTRUMENT AND MEASUREMENT

This chapter intend to; discuss in detail about each employability skills instruments previously mentioned in Chapter One, compares and obtain the list of employability skills considered necessary to graduates, discuss about three types of measurement method, namely, self-reported assessment, cognitive-oriented assessment, and SJT and finally giving an explanation on the steps of instrument development.

2.1 Introduction

Employability skills are generic skills that can be used in all types of jobs and work positions (Raybould & Sheedy, 2005) and these complement specific skills (Quek, 2005).

Employability skills are defined in several ways. These refer to teachable skills that are necessary in every level of job in organization across broad disciplines to secure and maintain jobs (Robinson, 2000; Singh & Singh, 2008). Curtis and McKenzie (2001) define employability skills as qualities of resourcefulness, adaptabilities and flexibilities possessed by employees needed to succeed in work and life as a whole. Quek (2005) defines employability skills as an ability of employees to draw upon the strength of their specialized know-how to resolve problems in workplace. The Ministry of Higher Education (MOHE) (2006b) in Malaysia views employability skills as craft skills that must be possessed by graduates in global employment. Singh and Singh (2008) define employability skills as the skills that graduates

possess in order to work. Based on these definitions, employability skills are generic skills that should be possessed by employees to perform.

This chapter is divided into six sections. The first section acts as an introduction. The second section describes the instruments of employability skills as mentioned in Chapter One. The third section examines the important employability skills required by graduates in their work which become the constructs of the instrument used in this study. Next, the fourth section discusses the three measurement methods i.e. self-reported assessment, cognitive-oriented assessment and situational judgment test (SJT). The focus of the discussion is on the strengths and weaknesses of each measurement method. Then, the best method was selected and used to develop the instrument. The fifth section describes the process involved in the instrument development. Finally, the sixth section provides a summary of this chapter.

2.2 Existing Employability Skills Instruments

As discussed earlier, there are five existing employability skills instruments. These are National Work Readiness Credential (NWRC), WorkKeys Career Readiness Credential (WorkKeys), Workforce Skills Certification System (WSCS), Workforce Alliance for Growth in the Economy Certificate Program (WAGE) (Alicea & Scott, 2007), and Malaysian Soft Skills Scale (My3S).

National Work Readiness Credential (NWRC) is one of the well-accepted work readiness instruments in the United States of America. It was developed by the National Work Readiness Council headquartered in Washington, DC. The instrument is targeted to the community of USA ranging from youth to adults

seeking entry level jobs. There are nine skills assessed in four modules. The first module, namely, situational judgment is a 60-minute assessment on the skills of cooperation, conflicts management and negotiation, critical observation, and problem solving. The second module is about reading with understanding, in which reading skill is assessed and the time dedicated for this module is 45 minutes. The third module i.e. "using math to solve problems" was developed to assess mathematical skills and the time given for this module is also 45 minutes. Finally, the oral language, which is the fourth module, was developed to assess listening and speaking skills. Thirty minutes are allocated for this module. The measurement methods used in the instrument are the cognitive-oriented assessment and the situational judgment test (SJT) (Alicea & Scott, 2007).

WorkKeys Career Readiness Credential (WorkKeys) is an online instrument developed by ACT, Inc., Iowa in the late 1980s (WorkKeys, 2012). It is being implemented in the United States of America, France and Italy (ACT, 2008). This instrument is used by employers to measure employee skills. It is also used by schools and colleges to prepare students for workplace. Specifically, this instrument is targeted to high school students and adults who seek for bottom or middle-level jobs (supervisory position). There are three assessments conducted in WorkKeys, namely, reading for information, applied mathematics and locating information. Respondents can also go for other six assessments of WorkKeys, namely, applied technology, writing, business writing, teamwork, observation, and listening. The instrument uses two measurement methods, namely, cognitive-oriented and self-reported assessment (Alicea & Scott, 2007; ACT, 2008).

Workforce Skills Certification System (WSCS) is an instrument developed under Comprehensive Adult Student Assessment Systems (CASAS) with cooperation between Sacramento-area schools and businesses (LEEDS) and Strumpf Associates (Eguez & Dayton, 2006). This instrument only focuses on five industries, namely, banking, construction, health, telecommunications, and high-tech. It targets youth and adults who are entering the workforce, work transition, incumbent workers, and dislocated workers in the United States of America. WSCS consists of certification assessment battery and certification assessment portfolio (Eguez & Dayton, 2006). Each component measures different skills. For certification assessment battery, the skills measured are reading, math, problem solving, and critical thinking. For the certification assessment portfolio, assessment is made on team skills, customer service, project development, and presentation. Additionally, there is an optional assessment that can be taken by respondents. The instrument is simulation-based and assesses communication, problem solving, critical thinking, teamwork, and customer service skill. The measurement method used is cognitive-oriented (Alicea & Scott, 2007; Eguez & Dayton, 2006).

Workforce Alliance for Growth in the Economy Certificate Program (WAGE) is a credential instrument that is currently used in Arkansas to help individual improve their basic employability skills. The target population is the same as NWRC, which is the community of the United States of America seeking entry level work. The instrument assesses four basic skills based on the Secretary's Commission on Achieving Necessary Skills (SCANS) research and Arkansas employer feedback (Alicea & Scott, 2007). These are math, communication, reading, and writing. Additionally, there are five optional tests that can be taken by respondents. These are

career scope assessment, WAGE post test, spatial relations test, dexterity test, and computer literacy test. The measurement method is cognitive-oriented ("Adult Education", 2011; Alicea & Scott, 2007).

Finally, the Malaysian Soft Skills Scale (My3S) is an online instrument under the auspices of the Ministry of Higher Education (MOHE), Malaysia. This instrument is used to assess the level of employability skills of higher education students of Malaysia. Additionally, it is used as one of the criteria in *Sistem Penilaian Institusi Pengajian Tinggi Malaysia* (SETARA). It was developed as a result of collaboration of researchers from seven public universities in Malaysia, namely, Universiti Putra Malaysia, Universiti Utara Malaysia, Universiti Pendidikan Sultan Idris, Universiti Teknologi Malaysia, Universiti Kebangsaan Malaysia, Universiti Islam Antarabangsa Malaysia, and Universiti Teknologi Mara ("Malaysian Soft Skills Scale", 2011). The target respondents for the instrument students of higher learning institutions in Malaysia. The instrument contains two-hundred and twenty items covering communication (54 items), critical thinking and problem-solving (30 items), teamwork (30 items), continuous learning and information management (30 items), entrepreneurship (20 items), professional ethics and moral (30 items), leadership (20 items) and self-appreciation (16 items) (MOHE, 2006b; "Malaysian Soft Skills Scale", 2011). Self-reported assessment is used in this instrument.

It might be helpful to adopt the existing instruments in this study, but there are three major issues in adapting these instruments: target population, employability skills, and measurement methods.

2.2.1 Target Population

The target population of the existing instruments can be categorised into two groups: (a) university students in Malaysia (for My3S), and (b) youth (with high school education) and adults. The existing instruments for this population are NWRC, WSCS, WAGE and WorkKeys. NWRC, WSCS and WAGE are appropriate for those seeking bottom-level jobs (non-professional positions), whereas WorkKeys is for those looking for bottom and middle-level of jobs (non-professional and supervisory positions).

On the other hand, the target population for new instrument in this study comprises graduates who are not covered in the existing instruments. Furthermore, the existing instruments do not assess skills needed in professional jobs that require graduates.

2.2.2 Employability Skills

In developing an instrument for this study, it is vital to incorporate the range of employability skills.

Each of the existing instruments assesses different skills. Eight skills are measured in NWRC and My3S, nine skills in WAGE and WorkKeys, and seven skills in WSCS. Even for the instruments that measure the same number of skills, such NWRC and My3S, the type of skills assessed are partially different. The skills measured by NWRC are oral language, reading, mathematics, ability to cooperate, conflicts resolution and negotiation, critical observation, problem solving and decision making, and learning responsibility, whereas for My3S, the skills measured are communication, critical thinking and problem-solving, teamwork, continuous

learning and information management, entrepreneurial skill, professional ethics and moral, leadership skill and self-appreciation. The variation in the number and type of skills makes it difficult for this study to determine what employability skills to measured. To address this issue, this study had to rely on previous studies (Section 2.3).

2.2.3 Measurement Methods

In the initial selection stage, work applicants are generally assessed on their competencies in their application. These competencies are obtained from self-reported assessment and cognitive-oriented assessment (Lievens et al., 2008).

Self-report (as used in WorkKeys and My3S) is a typical behavioural measurement method of data collection relying on respondents' subjective evaluation. This requires them to report honestly on their own behaviours or attitudes (Holbrook, 2008; Salkind, 2006). The score is then calculated based on a pre-determined scoring system and group norm (Urdan, 2005). Example of this method is when respondents have to declare any criminal offences they had committed in the past (Fielding, 2006). Generally, this type of measurement can be grouped into two broad categories, namely, unstructured open-ended question and structured closed-ended question. The main difference between the two categories lies in the answering format. In the open-ended question, respondents are to indicate their answers subjectively. In the latter, they provide categorical responses in the closed-ended question. The scale can be dichotomous (Yes-No), Likert or semantic differentials (Salkind, 2006). Unstructured open-ended assessment has the advantage of obtaining information that could not be captured using the structured closed-ended assessment.

It is as an optimal approach for qualitative data collection. But, it is ineffective for quantitative data collection. Quantitative data are best collected using structured closed-ended question. Self-reported measurement method is fast, cheap and practical, and is best used to gather information that cannot be obtained objectively (Salkind, 2006).

Despite its benefits, self-reported measurement method assumes that respondents are able to answer the questions posed to them and they are willing to do so. Unfortunately, the assumption is always unintentionally violated. This is because human beings have a limited and imperfect access to their internal mental processes and they tend to be uncomfortable with sensitive matters (Holbrook, 2008). In fact, they may not be able to give an accurate response to the question asked. This situation may also be triggered by the questions in context and the limitation on their memory. Furthermore, based on the theory of reward and punishment, which asserts that a favourable character will bring reward and minimize punishment, respondents may tend to build a favourable image of themselves. As a result, the assessment on their own abilities may be inaccurate (Denning et al., 1989; Holbrook, 2008).

Cognitive-oriented measurement method (as used in NWRC, WorkKeys, WSCS and WAGE) was established as a results of the criticisms made on self-reported assessment in the early days of psychological measurement, namely, introspection. In contrast to self-report assessment, cognitive-oriented measurement does not rely on respondents' self-evaluation. The measurement is based on observation and evaluation of respondents' ability of mind and behaviour (Salkind, 2006). This

method is said to be more capable of assessing the drive force and internal mechanisms behind the behaviour of respondents (Wagemans, 2004).

The psychological domains assessed by cognitive-oriented measurement method are problem solving, decision making, reasoning, memory, language, attention, performance, sensation, perception, emotion and motivation (Wagemans, 2004). In measuring cognitive skills, the items incorporated compensatory or conjunctive. It means that "the strength of one or some combination of cognitive skills is in compensations for the weakness in one or some of the remaining skills required in solving the item" (Leighton & Gierl, 2011, pp. 240). Hence, to obtain an accurate response on a particular item, respondents do not need to master all the measured skills. Depending on the domains that need to be measured, an item's response scale can vary. Like the self-reported assessment, an item's response scale of this method may be dichotomous or polytomous (e. g. Likert scale and multiple choice) (Leighton & Gierl, 2011).

In spite of the advantages, cognitive-oriented measurement has limitations. Even though it could validly predict the measured abilities, for example, academic performance, it has consistently shown a large adverse impact. This drawback has prompted employers to seek a more standardized test that is suitable for a diverse group of people (Lievens et al., 2008). In other words, this limitation has rendered the method inefficient.

In summary, because of the limitations of the existing instruments of My3S, NWRC, WSCS, WAGE and WorkKeys particularly on the three main issues discussed

above, it was decided that none of the instruments could be adapted in this study. Consequently, a new instrument should be developed from the beginning. The following section will discuss the selection of employability skills and measurement method adapted in this study.

2.3 Employability Skills of Graduates

Information about employability skills that are perceived to be important for graduates is obtained from past studies. These include works by Lowden et al. (2011), Saunders and Zuzel (2010), Singh and Singh (2008), Wickramasinghe and Perera (2010), Winstead et al. (2009), Wye and Lim (2009), and Yuzainee et al. (2011).

Singh and Singh (2008) conducted a study on what employers consider as employability skills among graduates. Employers were asked to rate the importance of each listed skill. The list of skills were adopted from previous studies. The employers were also asked about employability skills possessed by graduates working in their organizations. In addition, they were also asked about their preferences in hiring local and overseas graduates. The survey was carried out in Klang Valley using convenience sampling technique. The response rate for the survey was 53%. Based on the survey, the employability skills perceived to be important by employers are those related to communication, problem solving and adaptability, human relation, time management, English proficiency, information and communication technology (ICT), and leadership.

Wye and Lim (2009) investigated whether the employability skills possessed by graduates of private universities in Malaysia meet the needs of employers. The study also examined the effectiveness of personal quality and employability skills acquired at the universities concerned. The study involved 30 employers and 600 students from various private universities in Malaysia. Assessment was made on 20 skills based on the previous researches by Lees (2002) and The University of Western Australia [UWA] (1996). The result indicates that eight employability skills were perceived important by employers. These are critical analysis, decision making, oral communication, planning, problem solving, self-management, teamwork and written communication.

Winstead et al. (2009) examined curricular activities implemented in the business program at South Carolina State University, namely, Leadership and Professional Development Program (LAPD). The program is to enhance the employability skills of the university's graduates. In the program, students are exposed to practical knowledge, professional knowledge and knowledge based on experience to complement the academic knowledge. As a result of this program, graduates were found to show better performance in job interviews. The number of employers satisfied with the performance of graduates in their internship had increased by 40% and the graduates who were able to get jobs immediately after completing their studies had increased by 38%. The employability skills that are emphasized in the program are communication, interpersonal, teambuilding, and business skills.

Saunders and Zuzel (2010) sought to gauge the perspection of employers on important employability skills that graduates need to possess. A questionnaire listing

a total of 36 employability skills was developed and distributed to the selected employer-respondents. These employability skills were based on previous works by Murphy (2001), Tariq and Cochrane (2003), Tomkins (2004) and Yorke and Knight (2004). The employer-respondents were asked to rank the importance employability skills as they perceive. The results showed that the skills perceived as being important to graduates were enthusiasm/willingness to learn, questioning/listening, attention to detail, oral communication, dependability, integrity, commitment, cooperation, team-working, and understanding concepts.

Wickramasinghe and Perera (2010) explored the employability skills that employers consider valuable for graduates to have in the workplace in Sri Lanka, using self-administered survey on randomly selected 26 employers. The list of employability skills assessed was adopted from Coopers and Lybrand (1998), Cotton (1993), Dearing (1997), Knight and Yorke (2002a, 2002 b, 2003), Lees (2002) and Zinser (2003). The respondents were asked to indicate: (a) the level of importance of each listed employability skill when recruiting a graduate; (b) the level of employability skills expected of graduates; and (c) the measures taken to develop the employability skills of graduates who have been accepted. Results showed that the employability skills perceived as important are problem solving, attitude towards work, teamwork, learning skills and self-confidence.

Yuzainee et al. (2011) used the framework of engineering employability skills they developed earlier, Malaysian Engineering Employability Skills (MEES), to gauge the accuracy of the framework according to employers. The framework identified the employability skills engineering graduates in Malaysia need to have. The work

involved a questionnaire survey to 500 employer-respondents and found out that engineering graduates need to master communication skills, teamwork, professionalism, problem solving, and decision-making skills.

Lowden et al. (2011) conducted a study, commissioned by the Edge Foundation under the Scottish Council for Research in Education Centre, to explore the perceptions of employers on the status-quo of the employability skills of fresh graduates and the former's perception on what skills graduates need to have to enhance employability. Data collection was by survey of face-to-face interview carried out from 2009 to 2010. The results showed that employers consider team work, problem solving, self-management, business knowledge, literacy and numerical expediency, ICT knowledge, interpersonal and communication skills, ability to use own initiative to be valuable. So are the ability to follow instructions and leadership skills.

Table 2.1 provides a list of employability skills used by each study discussed above. A comparison was made on the employability skills perceived to be important to graduates. The importance of the various skills of each study concerned was ranked based on the frequency of their appearances in these studies. It is assumed skills that were highly ranked were important skill. Table 2.2 shows the ranking of employability skills in respect to the studies concerned. There were only three employability skills (ranked first and second) that were significantly important, and these are as follow:

Table 2.1 List of Employability Skills Used By Previous Studies

Singh and Singh (2008)	Wye and Lim (2009)	Winstead *et al.* (2009)	Saunders and Zuzel (2010)	Wickramasinghe and Perera (2010)	Yuzainee *et al.* (2011)	Lowden *et al.* (2011)
Communication skills	Oral communication	Communication	Oral communication		Communication skills	Good interpersonal and communication skills
Human skills	Written communication	Interpersonal	Enthusiasm	Self-confidence		Self-management
Problem solving	Self management	Business skills	Understanding concepts	Problem solving	Decision-making skills	Problem solving
Time management skills	Problem solving		Integrity		Problem solving	Knowledge of the business
Leadership skills	Decision making		Commitment	Positive attitude towards work	Professionalism	Initiative (but able to follow instructions)
Information and Communication Technology (ICT) skills	Critical analysis		Cooperation	Teamwork	Teamwork	Team working
English proficiency	Planning		Team-working			Leadership skills (where necessary)
	Teamwork	Teambuilding	Dependability			ICT knowledge
			Questioning/Listening	Learning skills		Literacy and numeracy relevant to the post
			Attention to detail			

1. Communication skills - Based on studies by Lowden et al. (2011), Saunders and Zuzel (2010) (oral communication), Singh and Singh (2008), Winstead et al. (2009), Wye and Lim (2009) (oral and written communication) and, Yuzainee et al. (2011).

2. Teamwork - Based on studies by Lowden et al. (2011), Saunders and Zuzel (2010) (dependability, cooperation and team-working), Wickramasinghe and Perera (2010), Winstead et al. (2009) (teambuilding), Wye and Lim (2009), and Yuzainee et al. (2011).

3. Problem solving - Based on studies by Lowden et al. (2011), Singh and Singh (2008), Wickramasinghe and Perera (2010), Wye and Lim (2009) and Yuzainee et al. (2011).

Most of the skills listed in Table 2.2 can be considered a subset of other skills. For example, the skills in Halpern (2003) and Schafersman (1991), critical analysis (Wye & Lim, 2009) and decision making (Wye & Lim, 2009; Yuzainee et al., 2011) may considered a subset of problem solving skill. Hence, the term "critical thinking in problem solving" is later used in this study to reflect problem solving skills. Additionally, Table 2.2 also shows the subsets of other employability skills, namely, personal quality, professional ethics and morality, and entrepreneurial skill. Firstly, the employability skills in ACNielsen (2000), Curtis and McKenzie (2001), Hodges and Burchell (2003), and Mathews and Redman (1996) proved that adaptability and human skills (Singh & Singh, 2008), self-management (Lowden et al., 2011; Wye & Lim, 2009), interpersonal skill (Winstead et al., 2009), enthusiasm (Saunders & Zuzel, 2010) and self-confidence (Wickramasinghe & Perera, 2010) may be

Table 2.2 The Ranks of Employability Skills Based on Previous Studies

Rank (No. of Studies)	Employability Skills	Rank (No. of Studies)	Employability Skills
1 (6 out of 7)	Communication skills	4 (1 out of 7)	Questioning/Listening
1 (6 out of 7)	Teamwork	4 (1 out of 7)	Attention to detail
2 (5 out of 7)	Problem solving	4 (1 out of 7)	Enthusiasm/Willingness to learn
3 (2 out of 7)	Leadership skills	4 (1 out of 7)	Integrity
3 (2 out of 7)	ICT skills	4 (1 out of 7)	Commitment
3 (2 out of 7)	Self-management	4 (1 out of 7)	Understanding concepts
3 (2 out of 7)	Decision making	4 (1 out of 7)	Self-confidence
3 (2 out of 7)	Interpersonal	4 (1 out of 7)	Positive attitude towards work
3 (2 out of 7)	Business skills	4 (1 out of 7)	Learning skills
4 (1 out of 7)	Planning	4 (1 out of 7)	Professionalism
4 (1 out of 7)	Critical analysis	4 (1 out of 7)	Literacy and numeracy relevant to the post
4 (1 out of 7)	Time management skills	4 (1 out of 7)	Initiative
4 (1 out of 7)	English proficiency		

considered subsets of personal quality. Secondly, those in Fallon et al. (2000), MacCann et al. (2009), Hashim (2009), and Mohamed et al. (2004) showed that planning skill (Wye & Lim, 2009) business skill (Winstead et al., 2009), enthusiasm/willingness to learn and understanding concept (Saunders & Zuzel,

2010), and knowledge of the business and initiative (Lowden et al., 2011) are subsets of entrepreneurial skill. Taking in account of the subsets, a new rank of employability skills can be made (Table 2.3).

Table 2.3 Ranks of Employability Skills by Considering Its Subset

Rank (No. of Studies)	Employability Skills
1 (6 out of 7)	Communication skills
1 (6 out of 7)	Teamwork
1 (6 out of 7)	Personal quality
2 (5 out of 7)	Critical thinking in problem solving
3 (4 out of 7)	Professional ethics and morality
3 (4 out of 7)	Entrepreneurial skill
4 (2 out of 7)	Leadership skills
4 (2 out of 7)	ICT skills
5 (1 out of 7)	English proficiency
5 (1 out of 7)	Questioning/Listening
5 (1 out of 7)	Learning skills
5 (1 out of 7)	Literacy and numeracy relevant to the post
5 (1 out of 7)	Attention to detail

The first six employability skills (ranked from first to third) are perceived as important for graduates as more than half of the studies had mentioned these. However, teamwork skill is inappropriate for this study because typically the screening process conducted by employers does not involve an assessment of teamwork skill. It will only be tested during the interview, which is normally

conducted after the screening process (Breaugh & Starke, 2000; Guion & Highhouse, 2006). Furthermore, there are organizations that carry out special group interviews to test the skills of a particular work group commissioned by the employer (Guion & Highhouse, 2006). Candidates who are successful in this interview are absorbed directly into the work group, whereas unsuccessful candidates will be reviewed by employers later to determine whether they could be accepted. This implies that teamwork skill is not an important skill during the early stage, and therefore is not appropriate during the screening process.

Consequently, five employability skills were selected to be incorporated in the instrument of this study. These are communication skill, personal quality, critical thinking in problem solving, professional ethics and morality, and entrepreneurial skill.

2.3.1 Communication Skill

According to Devito (2009), communication covers several forms: intrapersonal, interpersonal, interview, small group communication, communication in organization, public speaking and computer mediation.

Intrapersonal communication is a form of communication that allows one to communicate within oneself. Through this communication, one will be able to talk to oneself, learn about oneself and evaluate oneself. Interpersonal communication is a form of communication between two or more persons who are closely interdependent such as friends and family. Interview is a structured communication involving question and answer. Through this communication, one can learn about

other people. Small group communication is a form of communication in groups involving five to ten people. This form of communication takes into accounts the relationships and work requirements. Communication in organization is a form of communication among staff that occurs in the organization. This communication is about improving the formal and informal communication within the organization. Public speaking is a form of communication that involves speakers and listeners. A group of listeners may be small or large. The function of this communication is to provide information and to influence listeners. Finally, computer mediation is a form of communication between individuals or groups through the internet medium.

Next, communication principles should also be considered as they play an important role in shaping the communication and can be used to assist in improving the understanding of the communication process. According to Devito (2009), there are seven principles of communication. First, communication can be considered as a signal package. This is because communication involves a verbal and non-verbal package. Both of these packages always assist each other in facilitating mutual understanding in communication. For example, individuals who want to communicate something need to talk and make facial expressions and gestures to help listeners to better understand what it is that they are trying to convey. Second, communication is a process of correction. It can only occur between people who use the same communication system. For example, the communication will be problematic if the language used is different. Furthermore, the use and understanding of the word vary according to age, culture and country. In the real world, most people do not have the same communication system. Therefore, in every

communication, the correction process is always indirectly involved so that the message given is understood.

Third, apart from involving speakers and the listeners, communication involves a relationship between them. For example, the content of the messages conveyed by the employer to his subordinates may have similarities to the content of messages used by a friend-to-friends. However, when it comes to the differences in the relationships that exist between the two situations, the communication will be different. Fourth, communication is ambiguous. This is because each message can carry different meanings. Such ambiguity often occurs due to differences in the understanding of verbal and nonverbal communication used. Next, communication is a process of alternating, which means it has no beginning and end.

Sixth, communication between persons take into account the similarities and differences in response. This principle explains two conditions, namely, symmetrical relationships and complementary relationship. Symmetrical relationship shows the same attitude between the two parties. Conversely, in complementary relationship, different attitudes between the two parties are shown.

The final principle claims that communication cannot be avoided, backed up, and repeated. Specifically, communication will always occur intentionally or otherwise. Message that had been sent cannot be pulled back because it has reached the recipient, and it cannot be erased even it had been pulled back. Communication cannot be repeated because the first communication with the second one will not be

the same. Differences can arise from various aspects such as feelings of the recipient and the environment.

Therefore, in assessing communication, its forms and principles should be considered. However, for this study, the measurement of communication will only be focused on its important element. Unfortunately, this information cannot be identified through previous research because most of it did not discuss communication related to the extend on what is the most important element of communication. Hence, an alternative option that was used to obtain this information is by using an expert opinion.

According to Dr. Norhafezah, UUM, the overview of an individual's communication can be captured by looking at the level of his /her communication anxiety (CA). CA is a level of fear or anxiety related to communication with others. This view is consistent with previous studies that showed that a person with good communication skill should show a low level of CA and a weak communicator will demonstrate a high CA (McCroskey, 2009). According to McCroskey (2009), CA can be divided into four types, namely, Traitlike CA, Generalized-context CA, Person-Group CA, and Situational CA. Traitlike CA is associated with the type of communication possessed by the individual personalities in different contexts. Three contexts of this type of CA discussed in previous studies are oral communication, writing, and singing. According to theory of Traitlike CA, the CA level of each individual will often change throughout his/her life. Theory of Generalized-context CA states that individuals do not have the same level of CA in all contexts. For example, most individuals have a high level of anxiety when giving public speaking. However, in

cases of informal communication with friends, their anxiety level is low or possibly zero.

Person-Group CA describes the reaction of individuals in communicating with other individuals or group. For example, a teacher typically has a high level of CA when facing the headmaster. However, when confronting a student in class, his/her CA level is commonly low. The final type of CA is Situational CA. This type of CA looks at the reaction of individuals when communicating with other individuals or group in some specific situation. For instance, a student who makes an appointment with the teacher to ask questions may have low levels of CA. However, if he is asked to meet with the teacher in the office, the student's level of CA may be higher than the first situation.

As for the measurement in this study, CA was tested in general and not focusing on any type of CA describe above.

2.3.2 Personal Quality

Since the 19th century, employers had keen placing emphasis on the personal quality of prospective employees (ACNielsen, 2000; Curtis & McKenzie, 2001; Hodges & Burchell, 2003; Mathews & Redman, 1996). Hodges and Burchell (2003) and Spencer and Spencer (1993) assert that individuals who have good personal qualities appropriate to the position offered will be recruited sooner than those who only have the working skills for the post.

In general, personal qualities examined in this study refer to career-related qualities (CPQ). However, this does not mean that life personal qualities (LPQ) are different from CPQ. In fact, to be more accurate, CPQ is part of LPQ. As mentioned earlier, the Career Center of South Carolina University lists 121 types of LPQ. Some of the features in the list have been found to be in common with the CPQ in ACNielsen (2000), Curtis and McKenzie (2001), and Hodges and Burchell (2003).

There have been many studies on CPQ (ACNielsen, 2000; Curtis & McKenzie, 2001; Hodges & Burchell, 2003; Mathews & Redman, 1996). The important CPQ are as follows:

1. Enthusiasm

2. Integrity and honesty

3. Customer focussed

4. Resourcefulness

5. Flexibility and adaptability

6. Acceptance of feedback

7. Initiative

8. Capacity to learn

9. Self-control

10. Willingness to embrace change

11. Self-confidence

12. Independent problem solving

13. Interpersonal understanding

14. Practicality

15. Ability to benefit from on the job training

17. Innovativeness

18. Capacity to handle pressure

19. Commitment

20. Mature

21. Self-motivation

22. Personal presentation and grooming

23. Creativity

24. Self-esteem

25. Analytical

The list developed by the University of South Carolina shows that LPQ overlaps with other employability skills measured in this study. For example, appreciating feedback, innovative, creative, willing to take a risk, optimistic, and entrepreneurial are features measured under entrepreneurial skills. While conscientiousness, hard-working, industrious, and ethical are features measured in work ethics. Therefore, these features are suitable to be categorized under the CPQ. Mathews and Redman (1996) showed that communication is part of CPQ. This means that items developed for other employability skills in this study can also be used to measure CPQ.

However, there are also several other features of CPQ that do not overlap with other features of employability skills. The selection of CPQ for this study was based on the features considered important by employers. With reference to previous studies, there is a dearth of research on the importance of the various CPQ features, with only three studies used as references here. These are by ACNielsen (2000), and Hodges and Burchell (2003) and Mathews and Redman (1996).

Mathews and Redman (1996) compared the CPQs published in the 80s until the 90s. It showed that in ten years, CPQ that received attention from employers were enthusiasm and initiative. The findings are consistent with some of the results obtained by ACNielsen (2000). However, ACNielsen' study indicated that there are other CPQ that was more important than enthusiasm and initiative, namely, flexibility and adaptability, personal presentation, and maturity. As for Hodges and Burchell (2003), they showed that the most important CPQ is enthusiasm followed by flexibility and initiative. Hence, in measuring personal quality, the variables of enthusiasm and initiative was chosen with the assumption that demographic and geographic factors do not have a significant effect on the level of importance of CPQ.

According to Hodges and Burchell (2003), enthusiasm is the positive attitude shown by individuals in performing a task. Individuals who have a high level of enthusiasm will always try to solve any given task despite its difficulty. The feature will also help them to avoid despair. Initiative is defined as being proactive in initiating things.

2.3.3 Critical Thinking in Problem Solving

According to Halpern (2003) and Schafersman (1991), critical thinking is a process of descriptive scientific thinking to solve a problem and make a decision. Individuals with this skill are capable of gathering relevant and reliable information when making decision. An individual is said to think critically when the problem solving process involves the following (Schafersman, 1991, pp 56):

1. Asking questions that are appropriate to the problem.

2. Gathering relevant information.

3. Organize information obtained in a creative and effective manner.

4. Making a logical judgment based on the information obtained.

5. Coming up with a good conclusion about the problem and proposed solutions.

In the discussion between the researcher with Nordin Kardi (Psychology expert in UUM), generally, the existence of this skill in a person can be identified by looking at the response given when the person was asked to evaluate an issue. Individuals who respond descriptively by giving an explanation about the issue before making any decision is a person with critical thinking. However, if the individual does not give any explanation and directly jump toward the decision, he can be considered as not being critical in thinking. Therefore, to assess the skill, this study used the approach described by Nordin Kardi. In addition, the cognitive ability of graduates in analysing problems was also tested.

2.3.4 Professional Ethics and Morality

The assessment of Professional ethics and morality in this study refer to work ethics. If a person has work ethics, he/she does not need supervision to accomplish his/her job and does so sincerely (Huang & Cappelli, 2006). Work ethics is an important skill for graduates as it gives a direct impact on job performance (Boatwright & Slate, 2002). Ethics is essentially a moral knowledge that can lead to harmony (Skutch, 2004).

Work ethics has many dimensions (Geren, 2011; Miller et al., 2001). Miller et al. (2001) made a revision on work ethics from several previous studies including the research of Max Weber, a German scholar who developed theory of the modern work ethics. Miller et al. showed seven characteristics of work ethics:

1. Hard work

2. Self-reliance

3. Leisure

4. Centrality of work

5. Morality / ethics

6. Delay of gratification

7. Time appreciation

However, this study does not focus on the list because, according to Fallon et al. (2000) and Huang and Cappelli (2006), at the level of higher dimension, there are other factors that need to be addressed, which are more closely related to the assessment of the work ethics and some of the features listed by Miller et al. (2001). One of these factors is conscientiousness. This quality is one of five key factors listed in Big-Five personality characteristics. This theory is widely used in psychology to identify individual personality (Craighead & Nemeroff, 2004). Therefore, instead of the above list, this study focus on conscientiousness in making an assessment to professional ethics and morality.

According to Fallon et al. (2000) and Hogan and Ones (1997), conscientiousness consists of a combination of three characteristics i.e. willingness to work hard, orderliness, and self-control. MacCann et al. (2009) found that conscientiousness has

shown consistent characteristics. However, in their study, the terms "hard work" and "self-control" were replaced by Industriousness and Control, respectively. Even though there are three elements that make up conscientiousness, only industriousness is used to measure professional ethics and morality in this study. According to MacCann et al., if the nature of conscientiousness needs to be embedded into an individual in a short time, the focus of training should be given only to the modules on increasing the desire of the individual to work hard.

2.3.5 Entrepreneurial Skill

Originally, this skill is needed by entrepreneurs to conduct their business well, but it is not critical to non-entrepreneurs. However, over the years, employers have come to realize the importance of such skill to the success of organization. As a result, employers now consider this skill as being important (Deraman et al, 2010; Mohamed et al., 2004).

An entrepreneur was initially seen as an individual who ran a large-scale production. In the 17th century, entrepreneurship was considered unfavourably because of the losses as a result of a contract with the government at a fixed price. In the 19th century, the perception towards an entrepreneur changed favourably when it began to be regarded as a manager. In mid-20th century, an entrepreneur is portrayed as being secured for the future. In this age, entrepreneurs are seen as creative and innovative individuals who introduce new business dimensions such as new products, new distribution systems and development of new structures (Mohamed et al., 2004).

Mohamed et al. (2004) views an entrepreneur as "an individual who combines factors of production, introduce something new, identifying business opportunities, generates business ideas, innovates, and takes risk" (pp. 135). These characteristics are also perceived to be important not only to entrepreneurs, but also to individuals of other occupations. Many studies have focused on the identification of characteristics of successful entrepreneurs, and of the 27 studies from 1948 to 2006, there are over 20 entrepreneurial skills have been proposed (Hashim, 2009). However, only four characteristics have been repeatedly mentioned (Mohamed et al., 2004; Deraman et al., 2010), and these are incorporated into the instrument of this study:

1. The need for achievement – This consists of four elements, which are the desire to solve the problem independently, ensure that objectives are met, prefer to receive feedback, and tend to moderate risk.

2. The tendency to bear the risk – This is closely related to the need for achievement, but this feature focuses on a higher risk in exploring related activities and in financing.

3. Innovative – This characteristic refers to the ability to mobilize resources in a creative and innovative manner for business development.

4. Internal locus of control – This refers to the belief of an individual that he is able to control his future through his own efforts. This feature is closely related to the nature of optimism.

2.4 Selection of Measurement Method

As stated in Section 1.3, a good instrument should be efficient, realistic and standardized assessment (Lievens et al., 2008). On this score, this study had chosen

SJT as the measurement method that seems to fit such description. SJT shows the capability to fulfil the need of employers for an efficient (not based on the perspective of graduates/applicants), realistic (based on real work scenario) and standardized (has a low adverse impact) measure. Additionally, since SJT needs an implicit assessment through a problem situation, it will be difficult for respondents to make an attempt at falsifying their answers or to guess the best answers.

SJT is also known as professional dilemma, practical skill tests or situational test (Bizeur et al., 2010; Picard, 2007). Items developed using SJT measure reflect the form of situation or problems that appear in real work, and the options provided consist of different actions that can be taken by respondents in dealing with the situation. Therefore, in answering the item, the respondents are required to make judgments on the scenario given either in terms of problem solving or decision making (Bizeur et al., 2010; Lievens et al., 2008; Picard, 2008). As for the adverse impact, previous studies show that the differences in SJT's mean scores between races are much smaller than those of cognitive ability tests (Lievens et al., 2008).

SJT has generally a structure of multiple-choice in which that there is a question (problem situation) with options (actions). As the multiple-choice type, SJT can be used to measure various levels of difficulty and is suitable for all fields. The development process of the item calls for caution so that the response options listed will consist the best possible courses of action appropriate to the given situation. However, with the limited space in listing the response options compared to the diversity of action that can be taken by a person in a particular situation, some choice of action may not able to be listed. To overcome this problem, the respondents are

asked to respond only based on the provided list of actions. This will help avoid respondents from thinking too much on a given problem situation. A detailed discussion about this method is made later in chapter three. The next section explains the process involved in developing the measurement instrument. However, the focus of the discussion here is on the general process of instrument development. A detailed explanation of the development of the instrument is made in chapter four.

2.5 Instrument Development Process

The steps taken to develop the instrument are based on Chisholm et al. (2000), Dennis and Bocarnea (2005), DeVellis (2003), and Merten (2009). For example, Chisholm et al. (2000) aimed to develop an internationally acceptable instrument in collecting the data on the usage of service and other related characteristics of people with mental disorder. The target population were people who live in Amsterdam, Copenhagen, London, Santander and Verona. In developing the instrument, nine steps of development process were followed. These are as follows:

1. Identify the main desirable characteristics of an instrument in the respected area

2. Examine existing instruments.

3. Select an existing instrument for adaptation or develop a new instrument.

4. Translate into other languages.

5. Refine the content through focus group discussions.

6. Make a revision on the instrument.

7. Implement pilot test.

8. Re-examine the performance of the instrument (response rate, ease of completion).

9. Further revision before implementation.

Next, in *Scale Development* of DeVellis (2003), the eight steps process developing a scale of questions in an instrument are adopted and these are as follows:

1. Define the objective of instrument to be developed.

2. Create draft questions.

3. Select a common format and set of answer options for the questions.

4. Have experts review and revise the questions.

5. Consider inclusion of validation items.

6. Carry out pilot test.

7. Analyse the items.

8. Optimize the instrument.

The above steps were also adapted by Dennis and Bocarnea (2005) in their attempt to develop an instrument to measure leadership constructs based on Patterson's servant leadership theory. Merten (2009), in her book, *Research and Evaluation in Education and Psychology*, proposes six steps to be followed in developing an instrument. The steps were also adapted from DeVellis (2003) with some improvement. Several steps were merged to improve the efficiency in developing an instrument. These are:

1. Defining the objectives of the instrument.

2. Identification of the target respondents.

3. Revision on existing instruments.

4. Development of item.

5. Preparation of prototypes and conducting pilot studies.

6. Analysing and reviewing the measurement instrument.

According to Merten (2009), if problems of validity or reliability arise at the fifth step, a correction should be done, and pilot studies should be conducted again. This process will continue until the developed instrument is satisfactorily reliable and valid.

In essence, the steps in developing a data collection instrument are rather similar (Oppenheim, 2001; Merten, 2009) and no single framework of development is best. As the steps proposed in each study are similar, this study follows the six steps by Merten (2009) and uses the other proposed steps of instrument development for references. Figure 2.1 shows the steps used in this study. Details explanation on this figure was included in Appendix A.

The first step is to define the objectives of the instrument, clearly indicating the purpose of its development (Merten, 2009). In doing so, there are two considerations that need to be made. The first one is about the specific purpose of the proposed measurement instrument and the second one is about determining the information that need to be collected. Next is to identify the target respondents and the characteristic of the instrument (Dennis & Bocarnea, 2005; DeVellis, 2003; Merten, 2009). According to Merten (2009), "the relevance of the criterion of the instrument centers on the persons for whom the measurement is intended" (pp. 313).
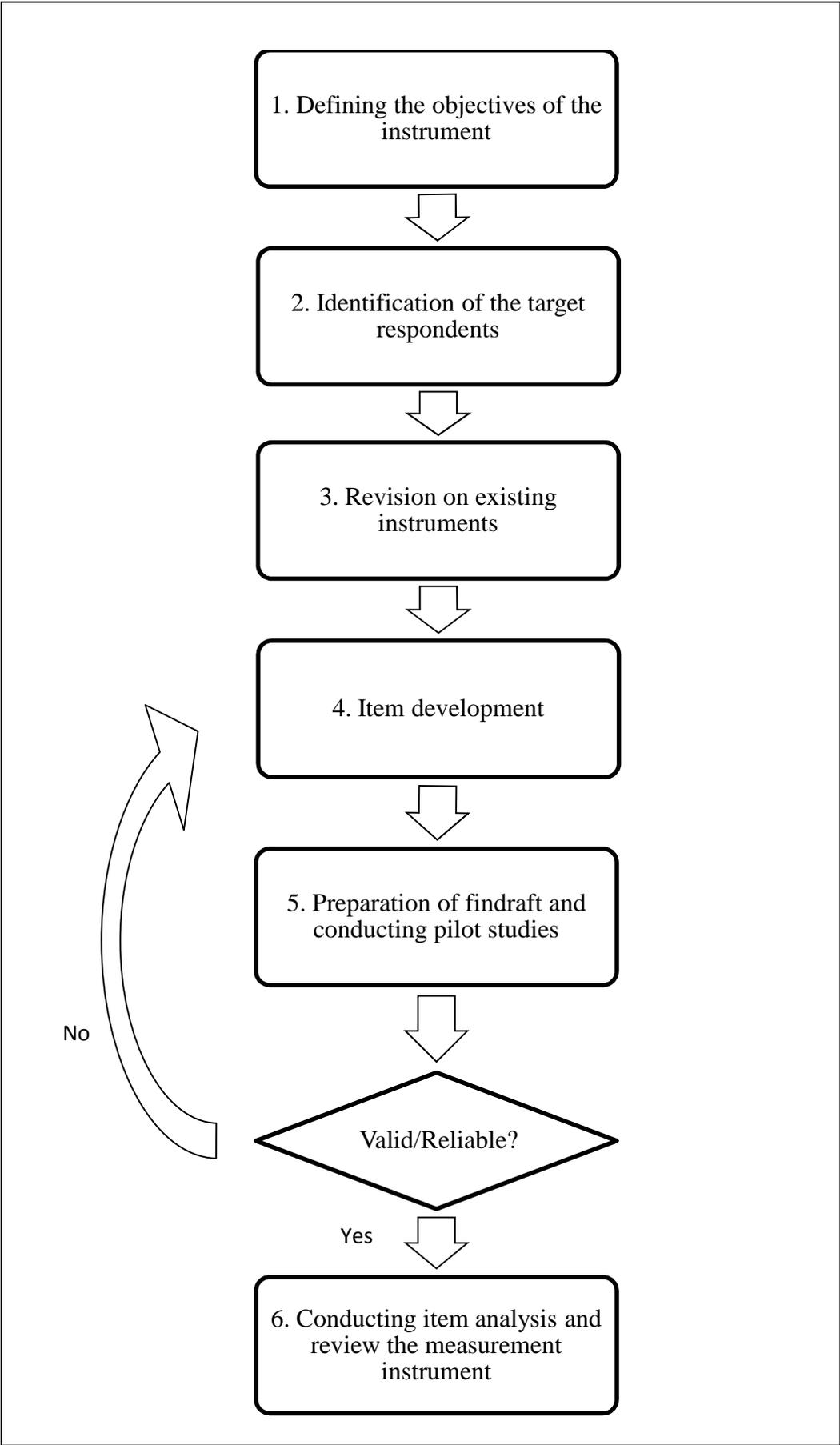
Figure 2.1: Instrument Development Process

The third step, if available, involves revision on the existing instruments (Chisholm et al., 2000; Dennis & Bocarnea, 2005; DeVellis, 2003; Merten, 2009). Next is the development of the draft instrument (Chisholm et al., 2000; Dennis & Bocarnea, 2005; DeVellis, 2003; Merten, 2009). This step is the most crucial part in the development process as this is where the instrument is constructed.

Once the instrument has been developed, the next step is to prepare the initial draft and the implementation of pilot studies. Subsequently, the validity and the reliability of the instrument can be determined (Chisholm et al., 2000; Dennis & Bocarnea, 2005; DeVellis, 2003; Merten, 2009).

The final step in developing an instrument is to conduct an item analysis and review the instrument (Chisholm et al., 2000; Dennis & Bocarnea, 2005; DeVellis, 2003; Merten, 2009). This step consists of an analysis using the data obtained at the previous step. Under this step, answers for each item will be reviewed in terms of their pattern to determine if the pattern suggests ambiguity or biasness.

2.6 Summary

The employability skills measured in the instrument consist of five skills: communication skill, personal quality, critical thinking in problem solving, professional ethics and morality, and entrepreneurial skill. As for the measurement method, situational judgment test (SJT) is used.

CHAPTER THREE

SITUATIONAL JUDGMENT TEST (SJT)

3.1 Introduction

As was discussed earlier, Situational Judgment Test (SJT) is considered the best measurement method to use in developing items for the instrument in this study. Compared to other methods, such as self-reported and cognitive-oriented assessment, SJT is more efficient, realistic and capable of standardizing across population (Lieven et al., 2008). Furthermore, many studies have shown the appropriateness of SJT in assessing work related employability skills of graduates (Motowidlo et al., 1990; McDaniel et al., 2001; Grubb; Nguyen et al., 2005; O'Connell et al., 2007; Rasmussen, 2009; Ployhart & MacKenzie Jr., 2010). This chapter describes SJT in more detail by considering its definitions, history, properties, development and strength.

3.2 Definitions of Situational Judgment Test (SJT)

Weekley and Jones (1999) and Bess (2001) describe situational judgment test (SJT) as a type of test that presents candidates with a hypothetical scenario based on realistic events with a number of response options. SJT is also known as written simulation tests (Joiner, 2008). According to Grubb (2005), Nguyen et al. (2005) and Rasmussen (2009), SJT is an assessment of work skill used to assess the capability of employee to deal with the situation on the problem that occurs in an employment setting. It is also regarded as a simulation method that requires respondents to make judgments in problematic work-related situations (O'Connell et al., 2007). More precisely, SJT is a measurement method of assessing opinions or interpretation of the

respondents on the given work situation, which reflects the actual working conditions (Legree & Psotka, 2006; Lievens & Coetsier, 2002; Garman et al., 2006; Weekley, 2010). One distinctive feature of SJT is its multidimensional nature of assessment. This method is capable of measuring multiple constructs per item measured, making it a good work-related measurement method. Most real situations are complex and involve multiple traits and abilities (Schmitt & Chan, 2006). Consider a situation, where a person is in conflict with his/her peer but needs to communicate with him/her. At first, the issue may be communication, but, actually, there are three skills to be tested, which are communication skills, personal quality, and professionalism.

3.3 History of Situational Judgment Test (SJT)

According to Weekley and Ployhart (2006), SJT-like item has been long used. However, there was an argument on what the SJT instrument was developed. Weekley and Ployhart (2006) note two different views with respect to the earliest existence of SJT. According to Whetzel and McDaniel (2009), the first instrument claimed to be the earliest instrument that applied the SJT item was used in the civil service examinations in the USA around 1873. Specifically, some items were tested by using a situational question. An example of the item is as followed:

> "A banking company asks protection for a certain device, as a trade-mark, which they propose to put upon their notes. What action would you take on the application?" (pp. 189)

In this view, it seems that the researchers' understanding on the characteristic of SJT was general and simple. They assume an item is consider as SJT if it is structured

with a situational question, even if there is no response option provided (Open-ended response). Actually, the format of open-ended responses make it more similar to the situational interview rather than SJT (Whetzel & McDaniel, 2009).

As for the second view, instead of just looking at the situational-based question, Additional feature of closed-ended response option was considered as another criteria for the SJT item (McDaniel et al., 2001). As the revision on previous research was made base on this view, they found that the first widely used SJT's instrument was the George Washington Social Intelligence Test that was developed in the early 20th century. Unfortunately, there are no other information given about that instrument.

Actually, the initiative of using SJT in the 20th century had been continued by other organizations afterward. SJT showed its importance in the World War Two in 1939 when used by psychologist of USA military in assessing the ability of the army on general knowledge, experience and level of rational thought. Then, in 1940, SJT was used to measure the potential of an individual as a supervisor. Examples of the instruments that use SJT are Practical Judgement Test, How Supervision?, Supervisory Practice Test, Business Judgement Test, Supervisory Judgement Test, and the Supervisory Inventory on Human Relations (McDaniel et al., 2001; Weekley & Ployhart, 2006; Whetzel & McDaniel, 2009).

SJT continued to gain its popularity in the late 1950s till early 1960s when it was used by large organizations in selection of employees. Examples of these instruments are the Early Identification of Management Potential and Management Judgement

Test. In the 1980s, SJT was used by the Office of Human Resource Management, USA, to assess the potential of job applicants at both officer and supervisor levels (McDaniel et al., 2001; Whetzel & McDaniel, 2009). As the usage widened, Motowidlo et al. (1990) studied the properties that should be considered for SJT instrument. Section 3.4 explains in detail these properties.

One of the common SJT instrument in the 21st century is the one developed by the National Work Readiness Council (NWRC), namely Work Readiness Credential (WRC) Situational Judgment Test. The instrument was developed to assess the employability skills possessed by job applicants. However, its targets are average entry-level workers (non-supervisory, non-professional) possessing relatively low literacy and numeracy skills only ("FAQs", 2011).

3.4 Properties of Situational Judgment Test (SJT)

According to Motowidlo *et al.* (1990) and Weekley and Polyhart (2006), the properties of an SJT item must consist of three characteristics. Those characteristics are 1) Items need to be present in a work-related situation, 2) Need to have a response option with the multiple-choice format, and 3) Scoring scheme need to be developed.

Generally, the first characteristic shows that SJT is a type of measure that suitable to be used in a job-related instrument as developed by this study. However, instead of the job-relatedness characteristic in SJT, the fidelity of the presented stimuli or the presentation of the situation in the instrument may be developed using a variety of formats. Typically, it can be presented through written form, video and multimedia

(McDaniel *et al.,* 2001). For the written form, situation is presented in a written format and respondent need to write down the answer based on the given response options. On the other hand, situation that was developed using the video and multimedia-based format consists of a video of scenario. In this format, commonly, question and response option were presented in visual and was assisted with narration (McDaniel *et al.,* 2001).

Additionally, the presentation for a set of SJT's instrument can be in more than one format. For example, the situation can be presented in a video format and the response option in a written format. However, it should be noted that the selection of the format will affect the accuracy level on the presentation of the item. For an instance, items presented using video was said to have higher accuracy than the written format. Still, video-based SJT has not been widely used because it involves high development costs and has low cognitive loading (McDaniel *et al.,* 2001; Olson-Buchanan & Drasgow, 2006; Weekley & Ployhart, 2006).

Next, for the second characteristic, the response given by the respondents must base on the given list of response option only. Again, similar to the situation presentation, response option can be presented in a written, video or multimedia format. However, additionally, in developing response option, the response style was the other things that also need to be considered. Generally, there are two components of response styles that need to be taken into account. The first component stress on the structure development of response option. Basically, the response option that is developed for each item is no relation to other items. However, depend on the item's difficulty level that wants to be developed, there are three types of response structure

available. Those structures are known as branched, nested and interactive (McDaniel *et al.,* 2001).

Branched item involve with the situation where the latter item that will be answered by the respondents is depending on the answers given by him in the former item (Weekley *et al.*, 2006). For the nested item, the developed item will consist of several questions that are based on a single scenario (Weekley *et al.*, 2006). Finally, as for the interactive item, it was known as the typical type of item for video and multimedia-based SJT. Respondents will be presented with a video or multimedia-based situation. Then, they will be led to differences video fragment depend on their response on the former item (McDaniel *et al.,* 2001). If the researcher wants to develop an easy SJT item, the basic structure of SJT item can be used. However, if they intend to develop a more difficult item, it can be developed by using among the three structures described above. However, it should be reminded that the difficulty of item cannot just be determined by looking at its structure. It would also depend on other things such as the subject that was being tested, its length and its fidelity. Hence, there still a possibility that an SJT item that was using the basic structure is difficult (McDaniel *et al.,* 2001).

The second component that should be considered in developing response option is related to the type of response that was expected to be given by the respondents. There are two types of responses that can be applied. Those responses are Knowledge response and Behavioural Tendency response (McDaniel *et al.,* 2001). Knowledge response tests the respondents on "what should be done" in a given situation. This type of response was said to be equivalent with cognitive abilities test

and job knowledge test. While for Behavioural Tendency response, it tests the attitude of the respondents or in other words, "what will be done" by them when faced with a given situation. It is equivalent with the personality inventory test. This type of response was also said to be easily manipulated by the respondents in obtaining good results (Hooper *et al.,* 2006; Weekley *et al.,* 2006).

The last properties that need to be considered are in terms of its scoring. For the subjective-based instruments (e.g.: essay, subjective question) that were developed to test an individual's ability, the score will be given based on the answer scheme and also involve a subjective assessment by the examiner. This means that in marking the assessment, the score was not necessarily based solely on the answer scheme. It was because, even though the answer given by the respondent was totally different from the scheme, if the answer was seems as acceptable to the examiner, then, mark will be given. This situation tell us that for such instrument, the scoring can only be made by an expert, which was the examiner. But, for the SJT instrument, expert judgment will only be involved in the development of answer scheme. After that, marking process can be made even by the non-expert individual as the scoring will be based just on the provided answer scheme. This characteristic is also being seen as equivalent to personality inventory test (McDaniel *et al.,* 2001).

3.5 Development of Situational Judgment Test (SJT)

The development of SJT item in this study follows that recommended of Motowidlo et al. (1990), and Weekley and Polyhart (2006), who propose the three properties of SJT. In developing SJT, each property was considered as a distinct step. With this consideration, the development of SJT in this study was divided into three stages

(development of situation, development of response option, development of scoring scheme).

3.5.1 Development of Situation (Question)

In developing work-related situations, several theories and models of employability skills were reviewed. This was to obtain a list of employability skills considered important for graduates. Next, after the employability skills has been identified, important elements to be measured at each of those skills were determined through discussions with experts in the related fields as well as references from previous studies. Then, using the information on several researches of workplace conflict (Kisamore *et al.*; Nussbaum & Johnson, 2010; Gupta *et al.*, 2011), situations that correspond to each related element was developed.

Kisamore et al. (2010) investigated the relationship between three types of social competencies, namely, political skill, self-monitoring and emotional intelligence, and the act of abusiveness and conflict at workplace. Three control variables were considered i.e. job satisfaction, number of hours worked, and gender. Results showed that workers in conflict have the tendency to be abusive against other workers. It also showed that individuals with high political skill and self-monitoring are more likely to demonstrate abusive behaviour when involved in workplace conflict. Nussbaum and Johnson (2009) propose a conceptual model of conflict management at workplace. One of the purpose of developing the model is to bridge two perspectives of conflict management behaviour i.e. conflict style and contingency approach. Gupta et al. (2011) attempted to synthesize the concepts,

definitions and styles of workplace conflict management and introduce a systematic approach to resolving conflict at workplace.

Additionally, expert opinion and situational questions developed by earlier studies (Bizeur et al., 2011; Picard, 2008; Harris, 2011; Wolff, 2011) were also being used to develop the situation in this study.

3.5.2 Development of Response Option (Answer)

Next, the development proceeded to the second stage. In this stage, the discussion focused on the matters related to the development of response options involving a number of issues such as resources used in developing the response options, its type, the level of difficulty and its level of delivering accuracy. Most of the SJT response options developed by previous studies refer on the opinions and recommendations by experts in the related fields (Weekley & Polyhart., 2006; Lievens *et al.,* 2008). As for this study, the development of response options was based on two sources, which are the expert opinion and the SJT workbooks by Picard (2008) and Bizeur *et al.* (2011) from France and United Kingdom, respectively.

The workbook written by Picard (2008) was a book that contains over 100 scenarios based on the SJT. It was published as a training source for General Practitioner test. As for the workbook written by Bizeur *et al.* (2011), it was a book designed to help in preparing candidates to take the SJT used by the European Personnel Selection Office (EPSO) for recruitment to the European institution. The usage of the workbook of Picard (2008) and Bizeur *et al.* (2011) was valid to be referred as the response options provided by their instrument seem to be similar with the work norm

of Malaysia. Even so, due to minor differences that appear in their employment setting, mainly in culture, some modifications in the situation and response were made so that it will be adaptable to the environment and working conditions in Malaysia.

As for the type of response option, this study tend to used the behavioural tendency response. Then, for the level of difficulty on the response option developed by this study, it can be classified as a moderate level response option. This is because, as the development of the situation, level of the difficulty of response option was assessed in terms of its length and the detail of description given (Motowidlo *et al.,* 1990). In terms of delivery, the method used was in written form so that will be consistent with the situation.

The list of response option developed by this study can be seen as effective as the method used in developing it was based on the articles that developed response option using empirical method (see Picard, 2008; Bizeur *et al.,* 2011). According to the study of Bergman *et al.* (2003), they have proven that choice of action resulting from empirical method was more closely related to performance measurement, cognitive ability and personality.

Next, after the response options have been developed, a discussion with the experts was made as to determined the answer scheme.

3.5.3 Development of Scoring Scheme

The last stage involves the development of scoring scheme. A detailed description of the scoring scheme for this study is discussed in chapter four.

Based on previous studies, two scoring method were available in the SJT. These are the forced-choice and continuous methods. In the simplest forced-choice method, there are three types of response formats: (1) The most appropriate action; (2) The most appropriate and inappropriate action; and (3) Sort the response options from the most appropriate action to the most inappropriate action.

The first format has only one answer that can be awarded with a mark. This is the common method used in a multiple-choice item. However, this scoring format has been modified to reduce the burden of the respondents in obtaining scores. For example, Hanson et al. (1999) set the scores in three levels as being " accurate answer," "less accurate answer", and "wrong answer." In the first answer, a full mark is given, while half mark is given for the second answer. No mark is given for a wrong answer. The scoring was developed based on the belief that less accurate response is not at all wrong. In circumstances where the right action cannot be taken due to certain constraints, the less accurate response is an acceptable alternative. Studies that have also used this modified scoring format are Hunter (2003), O'Connell et al. (2001), Ray et al. (2002), and Steven and Campion (1999).

The second format requires respondents to choose two answers. Motowidlo et al. (1990) developed the scoring based on five predetermined conditions (Table 3.1). In the first condition, "+2" will be given when the respondents give two answers correctly. In the second condition, they get to choose both answers correctly, but placed in reverse. For example, the most appropriate action is placed as the most inappropriate action and vice versa. For this condition, they will get "-2". Next, "+1" will be given when the respondent is only able to give one correct answer. After that,

the fourth condition is similar to the second condition. The difference is in the number of the correct answer chosen. For this condition, "-1" will be given for respondents that successfully choose only one correct answer, but is placed at the opposite position. Finally, if there is no answer that meets the four conditions, the respondents will be given zero. Calculation for this assessment is based on the summation of all marks obtained in each item.

Table 3.1 SJT's Scoring Scheme by Motowidlo et al. (1990)

Score	Condition
+2	Respondents managed to answer both situations correctly.
-2	Respondents correctly choose both answers, but have misplace them.
+1	Respondents successfully answered giving one correct answer.
-1	Respondents successfully choose a correct answer, but place in the opposite direction. Example: The action is best placed at the most inappropriate action or inappropriate action was placed as the most perfect action.
0	None of the above.

The last response format involves respondents having to look at all of the provided response options in the item and arrange them accordingly. Marks are given based on the arrangement made. The respondents must arrange five response options starting from the best response until the worst response (Weekley et al., 2006). An example of the implementation of this format can be seen in Weekly et al.'s (2004) and at the MCQBank website, developed to help medical student candidates in the United

Kingdom prepare themselves for Situational Judgment Test in the General Practice (GP) entry exam ("About MCQBank", 2012).

Weekley et al. (2004) use the Spearman's rank-order correlation to determine the scores of respondents. A high positive value indicates that response is consistent with the scoring scheme and a low positive value as inconsistent with the scheme; while a negative value indicates that the given response has an inverse rank-order of the scheme. In Situational Judgment Test (2012), the deductive scoring procedure is used. By default, each item in the instrument is worth 12 marks. Then, the mark is deducted depending on the accuracy of the rank-order response to the pre-determined rank-order in the scoring scheme. For example, assume that the scoring scheme for Item X is "A-B-C-D-E" and the response given is "B-A-C-D-E". Here, there has been a misplacement of the rank-order of "A" and "B". The "A" is one place away from the position that it was supposed to be. Therefore, one point is deducted from the total mark of 12. Because "B" is also at one place away from the correct order, another one point is deducted. In total, two marks are deducted. Hence, the total mark obtained for Item X is 10.

To choose between the three response formats that have been discussed above, there are several factors that need to be considered. Generally, each format can be considered as a different response format and can be selected according to the requirement of the study. However, in reality, each of these formats may also be regarded as a revolution to the previous format. For example, the second response format was developed to improve the first response format and the third format is an improvement made to the second format. The improvements made on the formats

were to address the criticisms labelled at previous formats. Weekley and Polyhart (2006) in their study made a comparison between the first and the second response formats, and they discovered that the second format is better than the first format in the level of variance. The usage of second format is seen to have improved the variance in the item scores and indirectly increase the validity of the instrument. Also, by using the format, the correlation value increases.

The third response format, i.e. sorting the response options from the most appropriate action to the most inappropriate action, was found to enhance the validity of performance predictions by 1.78 times better that of the first format and 1.5 times better that of the second format (Weekly, 2010; Weekley & Polyhart, 2006). However, the level of internal consistency for each response format is not significantly different. The problem of reliability in SJT is major and has been discussed ever since its development. The problem is believed to be caused by the multidimensional criteria of SJT. Specifically, any answer given by a respondent is based on a combination of abilities, experience and personality (Lievens et al., 2008). However, in terms of validity, SJT has an advantage. It is said that SJT is proven to be capable in representing real situation of work's problem. Hence, this criterion has made the SJT instrument an interesting predictive measurement instrument (Weekley, 2010).

The second scoring method of SJT is continuous method involving a Likert scale. This means that every response option of each item in the instrument has a score. For example, in the instrument developed by Chan and Schmitt (2002), the respondents were asked to rate the effectiveness of the response options provided under each item

on a six-point Likert scale. In scoring the instrument, each rating given by the respondents is compared with the rating given by the experts. Next, base on the degree of agreement by expert, a value of one, two or three is assigned. Finally, by using the total of the values, respondents' score can be obtained. Apart from that by Chan and Schmitt (2002), other studies have also used this method and have their own scoring procedures (e.g. Wagner, 1987; Wagner & Sternberg, 1985).

In sum, there are two types of SJT scoring methods: forced-choice method (Choose the most appropriate action, Choose the most appropriate and inappropriate action, Rank-order the response options from the most appropriate action to the most inappropriate action); and continuous method (Likert scale). In developing the instrument for this study, the forced-choice rank-order method (The most effective action to the ineffective action) was used.

3.6 Strength of Situational Judgment Test (SJT)

This section discusses the validity and reliability of SJT based on meta-analysis of 57 empirical studies of SJT published between 1990 and 2007 study (Lievens et al., 2008). The scope of advantages discussed were about SJT's criterion-related validity, incremental validity, construct-related validity, adverse impact, applicant perceptions, fakability, and susceptibility to practice and coaching effects and reliability.

3.6.1 Criterion-Related Validity

According to Lievens et al.'s (2008), SJT that was based on job analysis shows higher validity (0.39) than one that was not based on job analysis (0.29). It is evident

that SJT is an appropriate type of method for measuring work-related matters, including the measurement of employability skills. Theoretically, there are three assumptions that enable the SJT to measure employability skills: (a) that predictions for the future are based on past behaviour and events (the behavioural consistency principal); (b) that SJT can measure the intentions and goals of the respondents (goal-setting theory); and (c) that SJT is capable of measuring the variables that predict job performance.

With regards to item's presentation, video-based SJT was found to have the highest level of criterion-related validity (accurate in describing and delivering a situation). However, it may also can reduce the validity in the measurement. This is because it affects the measurement of cognitive abilities that require cognitive loading in reading components. Additionally, video-based SJT also may cause the situation of incident presented in the video to be misled. Respondents may consider the unintentional nuisance that appears in it as the main problem to that needs to be thought of and neglecting the main problem

3.6.2 Incremental Validity

Previous studies have been conducted to examine whether the use of SJT predict cognitive ability and personality well. Results showed that by using SJT, the validity of cognitive ability and personality test increased by 3% to 5% and 6% to 7%, respectively. These findings also indicated that SJT was able to increase the instrument's validity by about 1% to 2% better than that indicated by other measurement methods.

3.6.3 Construct-Related Validity

In the mid 1980s, SJT was said to be a suitable method for measuring tacit knowledge and practical intelligence, but was unsuitable for measuring cognitive ability. However, recent studies have refuted this opinion. Result showed that SJT has a good correlations with cognitive abilities, especially with the one that was related to employment issues. Specifically, the construct-related validity for SJT that is unrelated to job analysis was lower than the SJT related to job analysis with 0.30 and 0.50, respectively.

3.6.4 Adverse Impact

Even though SJT has only been tested on race and gender, generally, it has been argued that it did not have a problem of adverse impact problem. Previous studies have shown that differences in SJT's mean scores between races are much smaller than those of cognitive ability tests. For example, the standard deviation of the mean score between white and black respondents is 0.38. In terms of delivery accuracy, video-based SJT was found to have a less adverse impact because it has less cognitive loading.

With respect to gender, previous studies have found differences in the mean scores between men and women. The difference is 0.1 standard deviation in favour of women. This may be because of differences that exist in the natural response to the situations given in SJT. In addition, SJT item developed based on Behavioural Tendency response have shown a less adverse impact than that of instrument based on Knowledge response.

3.6.5 Perception of Applicant.

n general, respondents prefer to work one work-related instrument. Past studies have shown that respondents reacted positively to SJT. In terms of delivery format, video-based SJT give a more positive response than written SJT.

3.6.6 Fakability

According to Lievens et al.'s (2008), a study by Hooper et al. (2006) attempted to identify the fakability of SJT by assigning two groups with similar characteristics to answer the SJT instrument. The first group was required to provide the most honest response to each item and the other group to try falsify the answer. The results showed that the mean score for the two groups differs between 0.08 to 0.89 standard deviations. The study concluded that cheating in SJT is small.

3.6.7 Susceptibility to Practice and Coaching Effects

Past studies have shown that the scores obtained in the SJT can be improved through training. However, the increment was found to be small compared to other methods of measurement.

3.6.8 Reliability

Lievens et al.'s (2008) found that the internal consistency of SJT varied from 0.43 to 0.94 due to differences in the characteristics set in SJT. For example, in terms of sentence's length, longer sentences will raise its reliability. In terms of response format, referring to the section 3.5.3, the third format showed higher reliability (0.73) compared to the first and second formats with 0.60 and 0.24, respectively. However, due to the trait of multidimensionality in SJT, the usage of internal

consistency reliability in measuring SJT's reliability was questioned. Some previous studies stated that in measuring SJT's reliability, the suitable reliability measurement method to be used is the test-retest reliability method. The implementation of this method on SJT instrument shows that level of reliability obtained was between 0.77 and 0.89. The findings highlighted the fact that the SJT method has a good degree of reliability.

3.7 Summary

Based on the history of SJT that had been reviewed earlier, it can be concluded that SJT had been long used and the consistent expansion of its usage across time and fields proved that SJT was an efficient measurement method. Its level of efficiency may vary depending on the context. Despite this, SJT is said to be a better instrument because it is efficient, realistic and uniform (Lieven et al., 2008). In this study, the SJT items are developed using written format and do not involve complex structure such as branch and nested item format. Each item is presented with four ranked Behavioural Tendency response options.

CHAPTER FOUR
INSTRUMENT DEVELOPMENT AND IMPLEMENTATION

This chapter explains the methodologies applied in this study in achieving the objectives. Since there are three objectives, the main discussion in this chapter is divided into three parts. The first part explains the steps involved in developing a valid and reliable instrument. The second part elaborates the method used to ensure that the developed instrument is generic across graduates by taking into consideration their field of study. Finally, the third part outline the methods and process involved in obtaining the norm score distribution for the population of graduates.

4.1 Development of Graduate Employability Skills Instrument

As mentioned in chapter two, the development of the instrument in this study followed the six steps recommended by Chisholm et al. (2000), Dennis and Bocarnea (2005), DeVellis' (2003), and Merten (2009) as shown in Figure 4.1.

In the first step, the objectives of the instrument were defined. The second step involved identifying the target respondents to help the researcher decide on the type and difficulty level of the items. The third step was revision of the existing instruments of employability skills. The fourth step involved developing the items using situational judgment test (SJT). The fifth step was preparation of a draft instrument and implementation of pilot studies. Finally, an item analysis and review of the instrument was carried out. By using data of the pilot study, the final step was carried out to discover the difficulty level of each item perceived by the respondents
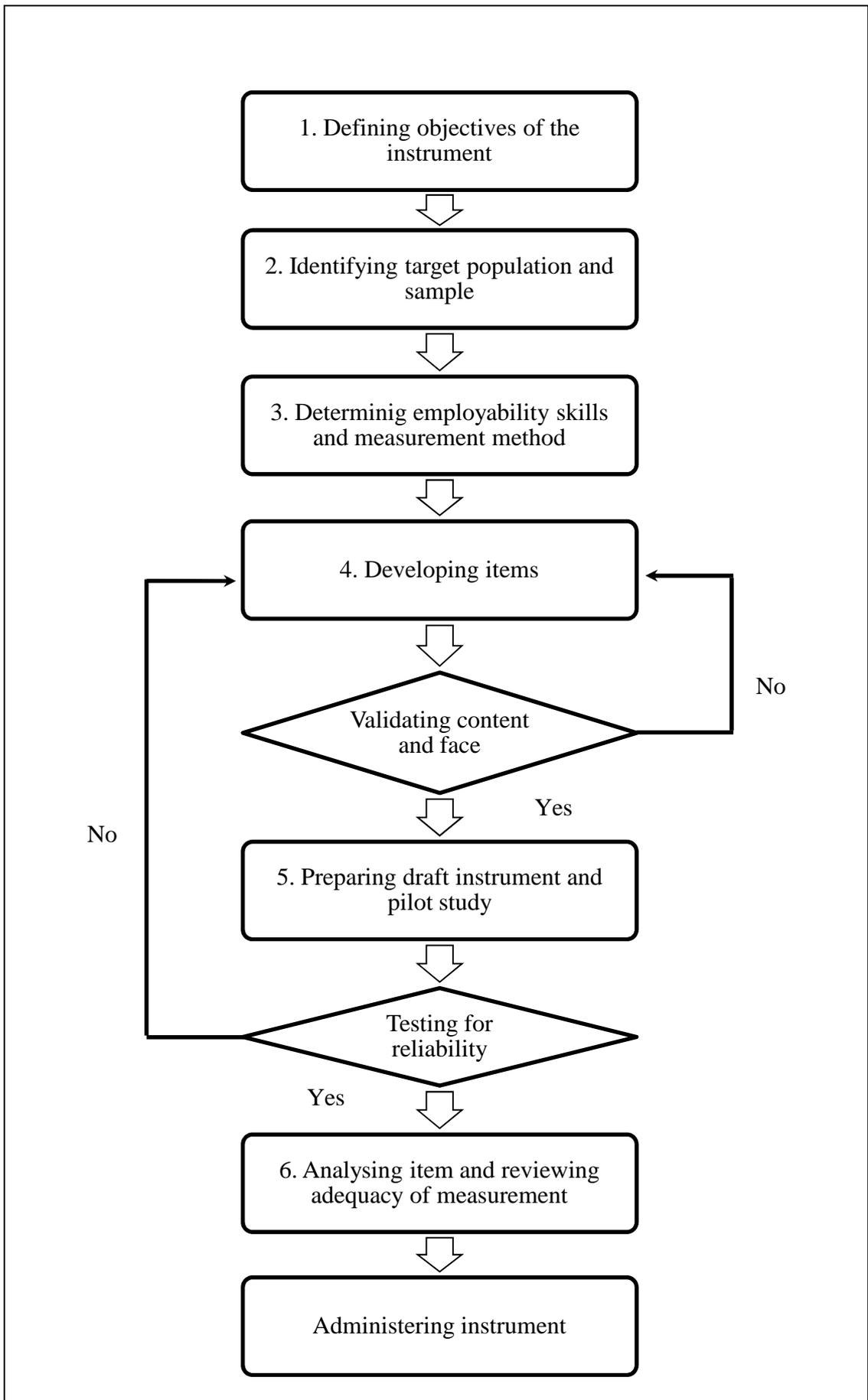
Figure 4.1 Flow Chart of Graduate Employability Skills Instrument Development

to ensure that the items are suitable to be used in assessing the employability skills of graduates.

4.1.1 Definition of the Instrument Objective

The objective of the instrument is to measure graduate employability skills using Situational Judgment Test (SJT). The situation was not just focus on working environment but other environment as well. The detail of the employability skills and SJT are explained in Section 4.1.3 and 4.1.4.

4.1.2 Identifying Target Population and Sample

The target population for this study only comprises of UUM graduates and the instrument is still at a prototype level. However, it was hard to obtain a fresh graduates without spending some time and money outside the university. As those two factors were the constrains for this study, the only chance that can be used in obtaining the target respondents was during their convocation. Unfortunately, the time for convocation can only be aligned with the time of the pilot study. Therefore, the implementation on the target population could only be administered in the pilot study. After that, the next group of graduates can only be obtained during the next convocation, possibly on another year. As it was time consuming, for the main study, it had been replaced with the final year students of UUM that was believed to possess similar criteria with UUM's graduate.

4.1.3 Determining the Employability Skills and Measurement Method

Studies by Lowden et al. (2011), Saunders and Zuzel (2010), Singh and Singh (2008), Wickramasinghe and Perera (2010), Winstead et al. (2009), Wye and Lim

(2009) and Yuzainee et al. (2011) were referred in determining the employability skills.

In summary, the employability skills suggested based on the analysis of the above studies shown five important skills for graduate. They are communication skill, personal quality, critical thinking in problem solving, professional ethics and morality, and entrepreneurial skill.

As for measurement method, SJT was chosen after reviewing the five existing instruments related to employability skills measurement. Based on the review, three methods could be identified i.e. self-reported assessment (WorkKeys and My3S), cognitive-oriented assessment (NWRC, WorkKeys, WSCS and WAGE), and SJT (NWRC)(Alicea & Scott, 2007). Based on the discussion on the strengths and weaknesses of each method in chapter two (Section 2.4), SJT was decided to be used.

4.1.4 Development of Items Using Situational Judgment Test (SJT)

Items developed in the instrument were based on the graduate employability skills and used the SJT method. At this stage, four expert opinions was solicited to achieve the content validity of the items. Appendix B shows the list of experts and Table 4.1 display a summary of the important elements for each tested skills as suggested by the experts.

Table 4.1 Measured Elements for Each Employability Skill

Employability skills	Measured elements
Communication skill	Communication anxiety
Professional ethics and morality	Willingness to work hard
Entrepreneurial skill	Need for achievement
	Tendency to take risk
	Innovative
	Locus of control
Personal quality	Need for achievement
	Flexibility
	Enthusiasm
	Locus of control
Critical thinking in problem solving	Descriptive ability
	Ability to think

Next, this study organized the number of items to be developed based on the number of elements that need to be measured in each tested skill. Table 4.2 shows the list of items along with the skills tested. The table shows that the measurement of Personal Quality involves more items than for other skills. The items were also not arranged systematically by cluster of skills; they were scattered randomly to avoid possibility of bias in case respondents could detect the pattern of arrangement and hence provide the answers accordingly.

Table 4.2 List of Items with Tested Elements

Item	Elements	Skills
1	Communication anxiety	Communication skill
2	Hard working	Professional ethics and morality Personal quality
3	Risk taking and need for achievement	Entrepreneurial skill Personal quality
4	Innovative and need for achievement	Entrepreneurial skill Personal quality
5	Flexibility	Personal quality
6	Thought	Critical thinking in problem solving
7	Communication anxiety	Communication skill
8	Need for achievement	Entrepreneurial skill Personal quality
9	Thought	Critical thinking in problem solving
10	Hard working	Professional ethics and morality Personal quality
11	Locus of control	Entrepreneurial skill Personal quality
12	Thought	Critical thinking in problem solving

Based on Table 4.2, the items were developed in three stages as recommended by Motowidlo et al.'s (1990) and Weekley and Ployhart's (2006). The stages are: (a) Development of the situation (Questions); (b) Development of response options (Options); and (c) Development of scoring scheme. The first two stages had been explained in detail in Chapter Three. Briefly, the development of the situation in the form of questions for each item was based on previous studies on conflict in the workplace (Gupta et al, 2011; Nussbaum & Johnson, 2009), expert opinion (Table 4.1) and questions regarding the situation provided in the workbooks and thesis of SJT (Bizeur et al, 2011; Picard, 2008; Wolff, 2011). Through these resources, situation (questions) has been developed. As for the response option, a forced-choice rank-order response format was developed based on the information obtained from the experts (Table 4.1) and the SJT workbooks (Bizeur et al., 2011; Picard, 2008). In each item, four options are available and each option is arranged from the most effective action to the ineffective action, in no particular order. The item is developed in the Malay language as the target respondents are Malaysian. Generally, the instrument is to test five skills (Communication skill, Critical thinking in problem solving, Entrepreneurial skill, Professional ethics and morality, and Personal quality) using twelve items. The following is a detailed discussion of each item.

Item One (Communication)

The first item is to measure communication anxiety of communication skill.

A friend from Penang has come to visit you in Kuala Lumpur. He uses the map to go to your place because he was not familiar with the roads there.

After work, you take her to go sightseeing in the city. While resting at a restaurant, an English tourist approached him asking for the way to KLCC. What is your action?

A. Leave it to your friend to help the tourists.

B. Provide navigation to the tourist and expect your friend to also help you by using his map.

C. Provide navigation to the tourist.

D. Provide navigation to your friend and ask her to tell them to the tourist.

The answer arrangement for this item is "C - B - D - A". "C" is the most effective action as it shows that the respondent is confident and has a low anxiety to communicate. "B" is selected as the effective action, which is the second best action because asking help from his/her friend shows that the respondent might not be comfortable in guiding the tourist because he/she is anxious. However, it is considered as an effective action because the objective in guiding the tourist is achieved. "D" is the less effective action because it shows that the respondent afraid to directly communicate with the tourist due to a high level of anxiety experienced by him/her. Finally, "A" is the most ineffective action because it shows that not only the respondent is afraid to communicate with the tourist but he/she is also afraid to be held responsible for the information that needs to be communicated.

Item Two (Professional Ethics and Morality/Personal Quality)

The second item is to measure the elements of hard working or enthusiasm that are important in professional ethics and morality, and personal quality.

Your group need to make a valuation report for the project that has been carried out. Each member gets the specific tasks to be completed. Three days before presentation, the report was compiled. You have found that the writing of several sections of the report is not as discussed previously. What is your action?

A. Help to identify the error and ask them to re-write the report in a day.

B. Ignore that parts and focus on the important part of the report.

C. Help other members in correcting that part and at the same time teach them how to write it correctly.

D. Request other members to make a correction on the respected part.

The answer arrangement for this item is "A - C - D - B". "A" is the most effective action because it shows the seriousness of the respondent in achieving the objective of his/her task without abandoning the sense of responsibility to other person. "C" is the second best action because the respondent shows that he/she is serious to finish the task. However, his/her motivation to do so is based on self-interest; he/she intends to complete the task even though by doing so no one else will gain any benefit. "D" is an ineffective action because it shows that the respondent lacks effort/motivation to achieve the objective of the task. Finally, "B" is the most ineffective action as the respondent totally abandons the task that needs to be completed and this action may affect the quality of the job later.

Item Three (Entrepreneurial skill/Personal quality)

The third item is to measure the element of risk taking, an important component in need for achievement. The assessment of this element is to capture the entrepreneurship and personal quality skills.

You have an excellent employment record with the company A. The original intention of your work is just to raise capital to start a business of Small and Medium Enterprises (SMEs). Now, you have enough capital. However, the salary and position in your career is very rewarding, secure, and you also get a promotion bid. What is your action?

A.   Take a six-month leave without pay to open and operate a business. Eventually, you will resign if the business is profitable.

B.   Continue working and at the same time open a business that does not interfere with your work at the office.

C.   In the near future, you will start a business as planned and resigned after the business is open.

D.   Forget the original intention because you have a job that can guarantee your future.

The answer arrangement for this item is "A - C - B - D". "A" is the most effective action to be taken as it shows that the respondent dares take the risk in starting a business without some backup finance for the next six months. This action also allows the respondent to focus on building his/her business. "C" is an effective action because this shows that the respondent dares take risk. However, this action is not the preferred one because the risk taken is too high and there may be no turning back

if the business failed. "B" is an ineffective action because it shows that the respondent has no gut in taking risk. However, in some sense, this action might be considered a brilliant move as it may help secure the finances until the business is stable. Finally, "A" is the most ineffective action as the respondent totally abandons his/her original objective. This action also shows that the respondent does not want to come out from his/her comfortable zone and take risk.

Item Four (Entrepreneurial skill/Personal quality)

The fourth item is to measure the element of innovativeness, which is also another component of "need of achievement" element.

You are a successful retailer. However, since the presence of a competitor in the neighbourhood, your profits have declined. What will you do?

A.   Strengthen the sales promotion at your store.

B.   Offer gifts and purchase coupons to lucky customers each week.

C.   Find a partner to raise capital to grow the business.

D.   Offer to the competitor to make a collaboration in the business.

The arrangement of the answer for this item is "B - A - C - D". "B" is the most effective action because it shows that the respondent is being innovative in handling the business competition. "A" is an effective action to be taken because it is the common strategy carried out by a business-minded person. Moreover, the promotion may involve some kind of innovativeness. "C" is an ineffective action because the respondent not only refuses to show innovativeness, he/she also does not want to take risk alone. Hence, by taking this action, the risk can be distributed to other stake

holder. Finally, "D" is the most ineffective action because it shows that despite being innovative, the respondent is not competitive.

Item Five (Personal quality)

The item is to measure the element of flexibility in the personal quality skill.

Your organization has sent you abroad for an external assignment for several years. That country is known to have a different culture and environment than Malaysia. What is your action in the first week there?

A.   Meet the staff who were at work during working hours.

B.   Making an observation for a few days before communicating with other staff.

C.   Interact with staff who are at work and try to understand the culture of the organization.

D.   Just focus on your work.

The answer arrangement for this item is "C - A - B - D". "C" is the most effective action as it shows the flexibility of the respondent and his/her eagerness to adapt in the new situation. "A" is an effective action because it also shows flexibility of the respondent in adapting with others. However, the pace of adaptability may become slower. "B" is an ineffective action because while time is being taken for observation, some "social wall" might have been indirectly built. Therefore, the process of adaptability may become harder. Finally, "D" is the most ineffective action as obviously, no effort is being shown to adapt with the new situation.

The sixth item is to measure the element of thinking in the skill of critical thinking in problem solving.

You want to buy a computer and have the necessary initial information. Having identified a number of options, you go to a computer store. You explain to the seller about the specs of computer that you need. He suggested a brand that is not on your list of choice. However, it is said to suit your needs and the price is cheap. What is your action?

A. Ask a few questions to the seller before making a purchase to ensure that the computer meets your need.

B. Re-evaluate your options and find an extra information on the recommended computer and other computer brands before making a purchase.

C. Accepted the proposal of the seller because he is more knowledgeable with computer.

D. Request permission from the seller to try out the computer so you will be more satisfied with any decision made.

The answer arrangement for this item is "B - A - D - C". "B" is the most effective action because as a critically minded person, once the decision has been disrupted, it is wise to take time and re-evaluate the available option. "A" is an effective action as it shows that the respondent does not directly accept the suggestion made by the seller. However, the usage of more sources other than the opinion of the seller is preferable. "D" is an ineffective action to be taken because, even though this action

is similar to "A", which is in the instantaneous conformation of the need, the source used in obtaining the clarification was inappropriate. Finally, "C" is the most ineffective action to be taken because obviously, there is no sense of critical thinking in the action.

Item Seven (Communication skill)

This item is to measure the element of anxiety of communication in communication skill.

You are the Chairman of a committee that is responsible for an opening ceremony for a program at your university. Before the program begins, you find that the emcee in charge was unable to attend due to health problems. The opening text and the committee's laptops is with him. There is no draft text that could be referred to and the time is running out. To ensure that the event is running smoothly, what should you do?

A. Only by using the program tentative, you become the emcee and conduct the event without text.

B. Provide a brief speech and take over the duties of emcee.

C. Instruct other committee members to take over the duties of emcee.

D. Instruct other committee members to take over the duties of emcee and you will be the last option to become the emcee if there is no volunteers.

The answer arrangement for this item is "A - B - D - C". "A" is the most effective action to be taken because it shows a high level of confidence in the respondent to be able to directly overtake the role without having any alternative and systematic

planning. "B" is an effective action because even though some planning needs to be made in advance, the capability to tolerate with urgency and take the role shows that the respondent has low anxiety. "D" is a less effective action because this preference shows that the respondent has quite a high level of anxiety in his/her communication. However, by allowing him/her self he to be last resort for the replacement, he/she can be considered as still having a bit of confidence. Finally, "C" is the most ineffective action because the run-away-like act shows that the respondent does not have the confidence to take on the position and is proof that he/she is uncomfortable to communicate in front of the audience.

Item Eight (Entrepreneurial skill/Personal quality)

This item is to measure the element of need for achievement that is important to entrepreneurial skill and personal quality.

> You have a restaurant. Foods provided by your restaurant always get compliments. One day, a group of customer criticizes that the food prepared by your restaurant has poor in taste, the way of serving the food is not right and they feels unsatisfactory. What is your reaction?
>
> A. Accept the criticism and will take an initiative to meet those expectation in the future.
>
> B. Just stay with your style as everyone has their own way of cooking.
>
> C. Accept the criticism and will consider it for the future.
>
> D. Accept the criticism, but you believe that every food served in your restaurant have the privilege of its own and should be maintained.

The answer arrangement for this item is "C - A - D - B". "C" is the most effective action because it shows that, despite of the current achievement, the respondent always gives him/herself a room for improvement so that better achievement can be made. "A" is an effective action because improvement for better achievement is still being made. However, the difference is that the thoroughness of the action taken may slow down the achievement process. "D" is an ineffective action because even though the criticism is accepted, no effort will be made to improve the menu. Finally, "B" is the most ineffective action to be taken as it shows no effort in obtaining better achievement.

Item Nine (Critical thinking in problem solving)

The item is to measure the element of thinking in the skill of Critical thinking in problem solving.

Nowadays, the source of petroleum is decreasing. Suppose that the Association of Consumer has recommend us to reduce the usage of petroleum to overcome this problem. What do you think?

A. Agree. Using of public services as a measure to reduce petroleum usage is cost-effective than the use of private vehicles.

B. Disagree. There are still lots of petroleum resources as many other potential areas have not yet being untapped.

C. Agree. The usage of petroleum contributed to pollution and cause global warming.

D. Agree. This could reduce the monopoly of oil and gas company and may lowering down the pressure on the market price of oil.

The answer arrangement for this item is "D - C - A - B". "D" is the most effective because this kind of thought can be considered as "out of box" and rarely thought by a common person as a consumer. Moreover, the agreement on the effect of monopoly on the respected company is relevant. "C" is an effective because the reason to agree for this action is a common consideration given the current environment condition. However, the reason might not be strong enough as there are other factors that cause pollution, such as the usage of air conditioning system and refrigerator. "A" is ineffective because even though the agreement is consistent with "D" and "C", the reason behind this agreement, which is cost saving, is not relevant. It is because the utilization of public transport may increase the daily cost rather than using a private transport. Moreover, by using public transport, the time spent to reach between places may increase. Finally, "B" is the most ineffective. Despite the disagreement, the reason shows that the respondent is simple-minded and selfish.

Item Ten (Professional ethics and morality/Personal quality)

This item is to measure the element of hard working or enthusiasm for the skills of professional ethics and morality and personal quality.

A staff have joined your team in a company. His presence is seen delay the work produced by your group. What is your action?

A. Make an overtime work so that any delay in your group work can be completed on time.

B. Ask the staff to expedite the process of his work.

C. Help the staff and remind him that you are always ready to help.

D.   Ask the staff to spend more time in the group so that he can be familiarised with the way the your group work.

The answer arrangement for this item is "C - D - B - A". "C" is the most effective action because it shows that the respondent makes effort not just in the given task, but also in helping others (especially new comers) in the adaptation process. By doing so, the new person will be able to pick up with the given task and reduce the problem of holding up the group. "D" is an effective action as it will have the same effect as "C". However, the action shows us that the new person needs to survive in the situation by him/herself. Without a mentor, the pace of adaptability might become slower. "B" is an ineffective action because by forcing, nothing good will come out. This is because, if the person is flexible, he/she might succeed in surviving in the group. However, if he/she has a fragile personality, he/she might end up as a failed person in the group. Finally, "A" is the most ineffective action because, even though this is an action of kindness, actually this action may give a negative effect on the new person. It is because as the task of the group has been completed, the person may end up learn nothing.

Item Eleven (Entrepreneurial skill/Personal quality)

This item is to measure the locus of control element that is important to entrepreneurial skill and personal quality.

You are offered with a higher post by the organization. The position is important and should be filled immediately. But you will be posting somewhere else. Working conditions for that position is in question as for the

past two years, two people in that position has end up resigning from the company. What is your action?

A. Reject the offer.

B. Accept the offer and consider it as a challenge.

C. Ask for a trial period.

D. Accept it as it is with the hope that everything will run smoothly

The answer arrangement for this item is "B - C - D - A". "B" is the most effective action as it shows that the respondent is confident with him/her self. "C" is effective because it also shows that the respondent is confident with him/herself. However, the request for a trial period shows that the respondent may feel curious and may have a low confidence level. "D" is an ineffective action because it shows that the respondent accepted the position with lack of confidence. However, by accepting the offer, he/she may learn something that may be able to build his/her the confidence level. Finally, "A" is the most ineffective action because it shows fear of the similar consequence that may befall him if he/she accepted the offer. He/she has no confidence in his/her working ability and assumes that the retirement was all caused by some unavoidable external factors.

Item Twelve (Critical thinking in problem solving)

This item is to measure the element of thinking in the skill of critical thinking in problem solving.

Your lecturer suggested that the number of person in a group assignment should be increased from three to five. What is your opinion on this proposal?

A. Agreed, as more ideas will be obtained during the process of task implementation.

B. Do not agree, because it may increased the time of group discussion.

C. Agreed, it will save time because the workload of each member will be smaller.

D. Do not agree, because the possibility to get passive members will be high.

The answer arrangement for this item is "A - C - D - B". In this item, the reason used for the action taken is important as it will become the objective of the discussion group. "A" is the most effective because the reason for this action is relevant and most beneficial compared to the reason for other actions. It also shows that the respondent seeks for the quality of the group rather that the quantity. "B" is an example of an answer that gives priority to the quantity rather than the quality. That is why it has been placed as the second best answer. "D" is ineffective as the concern of the respondent is irrelevant. Even though the concern is related to work quality, there is also possibility that the additional person is active. The main point is, chance should be given to this idea before any judgment can be made. Finally, "B" is the most ineffective because despite its disagreement, the argument is made by prioritizing the aspect of quantity.

A complete set of this instrument can be referred in Appendix C. Next, a focus group need to be conducted to obtain face validity. The group was consisting of five respondents between the ages of 24 years and 28 years. One of them was a newly-enrolled post-graduate student of UUM and the rest was UUM staff members who

were also graduates. The respondents were selected based on acquaintanceship because their level of employability skills was known. Before conducting the test, scoring scheme need to be developed. Next section discusses in detail the development of scoring scheme (stage three).

4.1.4.1 Development of Scoring Scheme

Several scoring schemes of SJT have been developed by previous studies such as those by Hunter (2003), MCQBank ("Situational Judgment Test", 2012), Motowidlo et al. (1990), O'Connell et al. (2001), Ray et al. (2002), Steven and Campion (1999), and Weekly et al. (2004). It was possible to find a lot of different scoring schemes in SJT as there is no absolute rule of scoring. Weekley et al. (2006) propose that a systematic research is needed to examine the scoring of SJT. There has been no research to determine the best scoring scheme to be used in SJT. This has allowed this study to develop a scoring scheme for the instrument. By modifying the SJT scoring method described by Motowidlo et al. (1990), this study has developed a scoring method for this study.

In developing the scoring scheme for this study, the response order and categories were identified in advance. Table 4.3 shows the four levels of response order in two categories used in this study. The first and the second level of response order were categorized as "Effective Action," and the third together and the fourth as an

Table 4.3 Levels of Response Order and Categories

Order	Response	Category
1st	The most effective	Effective Action
2nd	Effective	
3rd	Not effective	Ineffective Action
4th	Most ineffective	

"Ineffective Action." These response order and categories should be identified prior to the implementation as they would be taken into account in determining the scores of respondents later. Specifically, the score was based on the combination of response order by the respondents. Table 4.4 summarise the characteristic of response available in the instrument with their scores.

Table 4.4 Characteristics of Response and Scoring

Response Characteristics	Score
There is no answer sequence placed under the correct category	0
No order of correct answer but one effective action and one ineffective Action is still compiled under the right category	1
Only one correct arrangement of ineffective action	2
Only one correct arrangement of effective action	3
Only one correct arrangement for each effective and ineffective action	4
There is no right order, but each answer was still in the right category	5
Only two correct arrangement of ineffective action	6
Only two correct arrangement of effective action	7
Perfect answer	8

If there is no correct order of responses, the respondents are likely to obtain a score of zero, one or five points depending on the category in which the order was made. The respondents will be given zero mark if the responses were not placed in the right category. This is the lowest mark that can be obtained by the respondents for each item. One mark will be given if there is no accurate response arrangement, but with a condition that one effective action and one ineffective action are arranged in the right category. This is because even though there is no precise answer, the respondent would still be able to determine some of the actions alleged effective and not effective. Five points will be given when there is no right answer, but the order of each response is still in the appropriate categories. This is because in contrast to the respondent who obtained one mark, although there is no right answer, the respondent is still able to determine the categories of action that is effective and not effective as a whole.

Next, if the respondents succeeded in placing one response in the accurate order, they have the potential of getting two or three marks. They will be given two marks if one response for the ineffective action is accurate. Three marks will be awarded when they succeeded in providing one correct response for the effective action. Even though both conditions indicate that there is only one accurate answer in the arrangement, more points are awarded to the respondents who gave the right order for the effective action category because, in general, individuals who know how to act effectively are better than those who only know how to avoid ineffective act, without knowing the effective action to take.

Respondents who managed to arrange the correct order of two responses in one item have a chance of getting four, six, or seven points. Four points are given when they succeeded in arranging accurately one response of effective action and one response of ineffective action. Although this condition is almost the same as for those respondents awarded with one point, significant differences were indicated by the respondents given four points because there is the right order of answers in both categories.

Respondents will be awarded six points when there are only two accurate orders of responses given by them in the ineffective action category. In contrast, seven points are given when respondents correctly arrange the order of two responses in the effective action category. In this case, the situation shown is similar to the respondents who obtained two and three marks. The difference is just in the number of answers correctly ordered.

Respondents who managed to give the exact order of responses as in the answer scheme will be given eight points, the highest score that can be obtained for each item.

Since each item is provided with four response options (A, B, C, and D), overall, there are 24 combinations of the response order that can be formed from each item. Table 4.5 shows the whole combination of responses and marks that can be obtained by the instrument if the answer scheme is "A - B - C - D". This suggests that part of the response characteristics described above can be used to more than one answer combination pattern. For example, if the answer scheme for item X is "A - B - C -

D", based on Table 4.5, there are four combinations of answers that can result the respondents in obtaining a zero score. Those answers are "C - D - A - B", "C - D - B - A", " D - C - A - B" and "D - C - B - A".

Table 4.5 Examples of Response Combination and Scores for the Answer Scheme "A - B - C - D"

No.	Response	Score	No.	Response	Score
1	A - B - C - D	8	13	B - C - A - D	2
2	A - B - D - C	7	14	B - D - C - A	2
3	B - A - C - D	6	15	C - A - B - D	2
4	B - A - D - C	5	16	D - A - C - B	2
5	A - C - B - D	4	17	B - C - A - D	1
6	A - D - C - B	4	18	B - D - A - C	1
7	C - B - A - D	4	19	C - A - D - B	1
8	D - B - C - A	4	20	D - A - B - C	1
9	A - C - D - B	3	21	C - D - A - B	0
10	A - D - B - C	3	22	C - D - B - A	0
11	C - B - D - A	3	23	D - C - A - B	0
12	D - B - A - C	3	24	D - C - B - A	0

After scoring on each item is done, the average score of the items that measure the same employability skills are calculated. Then, based on the average score, respondents' employability skills are assessed. Table 4.6 shows the guidelines used in this study to assess the employability skills possessed by respondents.

Table 4.6 Guidelines in Assessing the Level of Respondent's Employability Skills

Score	Employability Skills' Level
0 - 2.66	Low
2.67 - 5.33	Moderate
5.34 - 8	High

Prior to the fifth step, which is the preparation of draft instrument and pilot study, a face validity test was carried out. This is to ensure that the measurement of the developed instrument is consistent with the objectives of the study. All of the respondents were given the initial draft of instrument. The draft has two parts: Part A, Situational Judgment Test, and Part B, Respondent Information. Part A has 12 items measuring five employability skills. Answers for each item were to be given in the space provided next to the response option. Part B requires the respondents to fill in their demographic information and their opinion about the instrument. To further strengthen the face validity, they were interviewed to provide a further feedback.

4.1.5 Preparation of Draft Instruments and Pilot Study

Once the content validity and face validity had been ascertained, a pilot study was carried out. The targeted respondents were the 592 graduates who attended the 23rd Convocation Ceremony (Second Session) of UUM at the Muadzam Shah Hall (DMAS), UUM, on March 19, 2011. The target population for the pilot study was quite small because the ceremony was a continuation of an earlier convocation that was adjourned due to flood. In determining the sample size of this population, simple random sampling was used. The formula used to calculate the sample size is shown in Equation 4.1 (Kalton, 1983).

$$n = \frac{z^2(\sigma^2)\left(\frac{N}{N-1}\right)}{ME^2 + \left(\frac{z^2(\sigma^2)}{N-1}\right)} \qquad (4.1)$$

Where

n   = Sample size

N   = Population size

z   = at the 95% confidence level

σ^2 = Variance

ME   = Margin of error

The following is the calculation made to obtain a minimum sample size that was considered in the pilot study.

1. Find the estimate of σ^2, the population variance,

$$s \approx \frac{8-0}{6} = \frac{8}{6} = 1.333$$

$$s^2 \approx (1.333)^2 = 1.778$$

2. With a 95% confidence level ($z = 1.96$) and a margin of error (ME)= 0.1 (10% from mean), the sample size are,

$$n = \frac{1.96^2(1.778^2)\left(\frac{592}{592-1}\right)}{0.01^2 + \left(\frac{1.96^2(1.778^2)}{592-1}\right)}$$

$$= 317.36 \approx 318.$$

So, to make sure that the sample is representative of the population, at minimum, a sample of 318 graduates were chosen.

96

4.1.5.1 Reliability Test of the Draft Instrument

As mentioned earlier in chapter three, SJT method has a multidimensional characteristic and this results in a low Cronbach's alpha value (Weekley, 2010; Whetzel & McDaniel, 2009). In other word, SJT is not suitable to be tested across item for reliability. To solve this problem, the measurement of variables should be made separately (unidimensional). However, it will only be effective if each variable is measured by numerous items. If the number of items for the variable is small, then it can also contribute to a low level of reliability (Schmitt, 2006).

The most suitable reliability test in measuring the reliability of SJT instrument is the test-retest reliability (Weekley et al., 2006) as it measures reliability in terms of its repeatability (stability) and not internal consistency (Zikmund, 2003). Even so, the method was unable to be used because the data obtained in this study did not meet the requirements for the method to be used. Specifically, it has no time series data.

So, since the reliability of the instrument cannot be proved directly through Cronbach's alpha, this study uses alternative approaches. However, these approaches require the data of the main study. The approach meant by this study was by comparing the scores of pilot and main study as to looked for instrument's stability. So, at this stage, the reliability test was put on hold until the data of the main study was available.

In addition, as the alternative approach in finding reliability involved the analysis on each item in the pilot and main study, it has automatically incorporated the sixth step

of instrument development, which was an item analysis. Therefore, it was assumed that the sixth step has been done.

4.3 Instrument Administration

4.3.1 Sampling Method

Next, the implementation of the instrument for the main study was using the sampling frame of respondents that composed of the final year local students from various undergraduate programmes at UUM. The target population of the main study was 10161 respondents. Seven thousand two hundred and sixty seven were from qualitative programmes and 2894 were from quantitative programmes. Shown below are steps in obtaining sample size:

1. Based on the pilot study, the estimate of σ^2, the population variance is

$$s^2 \approx (1.333)^2 = 1.778$$

2. The value of z and margin of error used in the main study are the same as used in the pilot study (z = 1.96 and ME = 0.1). The calculation is as follows:

$$n = \frac{1.96^2(1.778^2)\left(\frac{10161}{10161-1}\right)}{0.01^2+\left(\frac{1.96^2(1.778^2)}{10161-1}\right)}$$

$$= 639.97 \approx 640.$$

The calculation indicated that the minimum sample size required was 640. But, 1200 respondents were selected for the sample and since there were two major fields of study (Qualitative and Quantitative), 600 respondents were taken from each. The detail of the selections was explained in Section 4.3.1 in Chapter Four. Basically there are 21 programmes under qualitative study, and 10 programmes under

quantitative study. 24 students were randomly selected from each programme under qualitative study and 50 students from each programme under the quantitative study.

The population was first stratified based on qualitative and quantitative fields of study. Then, both strata were further stratified into undergraduate programmes (Figure 4.2). A simple random sampling was employed to choose the programmes. After the final list of programmes had been produced, copies of the instrument were then distributed. Three days were allocated to the respondents to give their responses. Next, the data obtained from the distributed copies of instrument were transferred into a spreadsheet (Microsoft Office Excel).
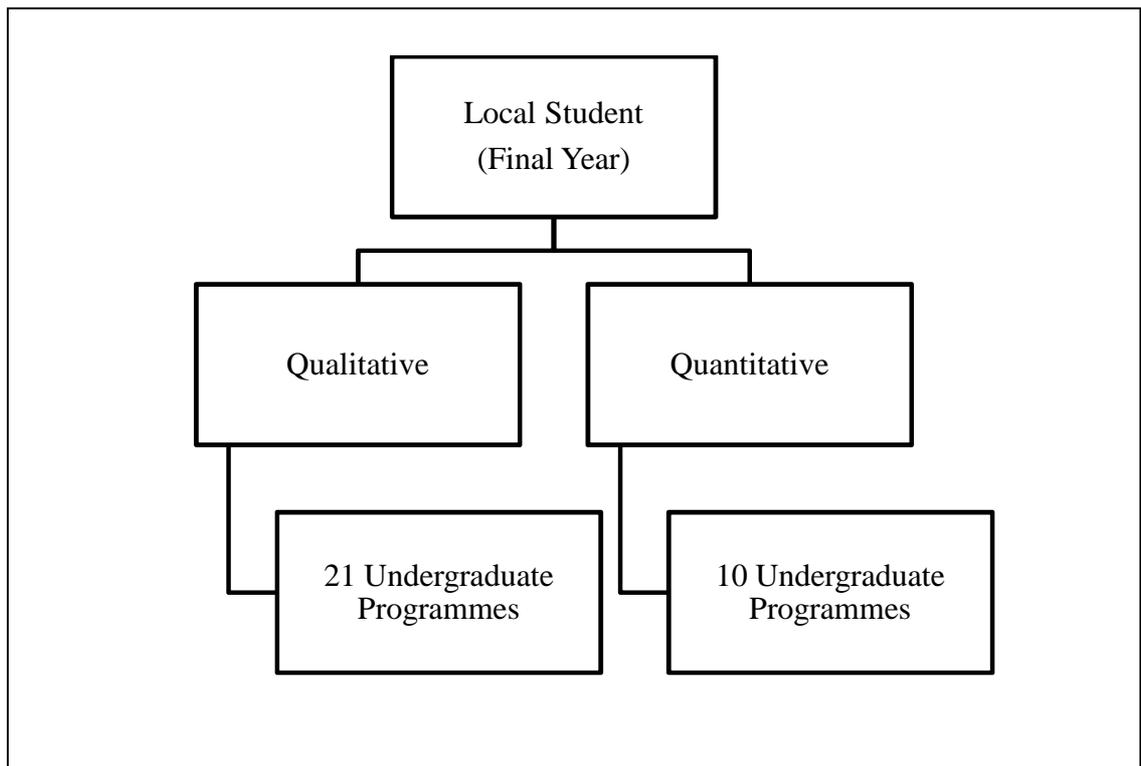
Figure 4.2 Flow Chart of Sample Selection

Then, these were checked for any errors before the scoring process. After the scoring process was completed, as the data for the main study has available, reliability analysis that was put on hold in Section 4.1.5.1 can be proceed.

Firstly, the data of the pilot and the main study were observed using boxplot, Cronbach's Alpha and exploratory factor analysis (EFA). Those were conducted as to explore and get an overview about the consistency among items within the same skill and across studies. After that, the process continued by exploring the frequency of scores for both studies in detailed as to support on any argument made during the previous observation. Finally, as to confirm the similarity of scores between pilot and main study, Chi-Square test for homogeneity was conducted. The following is the hypothesis used in conducting Chi-Square test for homogeneity:

$H_{0:}$ The scores of for each item pilot and main study are homogenous

$H_{1:}$ The scores of for each item pilot and main study are not homogenous

Then, hypothesis testing was conducted as to achieve the second objective, which is to ensure that the developed instrument is capable of measuring employability skills of graduates across the quantitative and qualitative programmes.

4.3.2 Obtaining Generic Instrument

The following was the hypothesis used in this testing :

$H_{0:}$ There is no difference in scores between respondents with quantitative and qualitative background studies.

H $_{1:}$ There are differences in scores between respondents with quantitative and qualitative background studies.

As mentioned at the end of the previous section, this hypothesis was tested to prove that the instrument of this study can be used by all graduates regardless of their undergraduate programmes (the second objective of the study). The independent sample t-test need to be used. However, this test can only be used if the scores satisfy the assumption of normality. Alternatively, Mann-Whitney U test will be used (Coakes & Steed, 2007).

Next, this study calculated the norm score distribution of the respondents. This was done to achieve the third objective of the study.

4.3.3 Norm Score Distribution

The use of norm score allows comparisons to be made between individuals in the group that share the same criteria (Salkind, 2006; Cohen & Swerdlik, 2009). However, this score cannot serve as a benchmark to measure whether the person has a good level of skill or not. If there are many respondents who scored low, the norm score would also be low. Thus, even though there are individuals who manage to score higher than the norm score, they still cannot be categorized as having a good level of skills because although they have exceeded the prescribed norms score, it may due to a low benchmark of the norm score (Guion & Highhouse, 2006).

For example, a norm score obtained through the implementation of the instrument can be used to make comparisons of the employability skills levels possessed by all

final year local students at UUM. Individuals who score close to the norm score are said to have a standard level of skills within the population tested. Students who scored above the norm score may be said to have better employability skills than most of other final year local students. Conversely, those who scored lower than the norm score are said not to have a level of employability skills as they should have been. Some examples of a norm score that can be used are Percentiles, Stanines and standard scores of Z-Score and T -Score (Salkind, 2006).

In order to obtain a good norm score distribution, a researcher needs to repeatedly conduct studies to several groups of samples taken from the same population (Berger, 2006; Good, 2005). However, due to various constraints, it is almost impossible to obtain samples repeatedly from the same population. Even if it is possible, it is better to take a one-time large size of samples. This is because estimation made using large samples is more accurate (Good, 2005). Therefore, an alternative for this problem is by using resampling method. These is generally seen as a modern method capable in replacing the parametric method. It has the strength to carry out parametric procedures without the need to meet any parametric assumptions (About Resampling, 2012; Berger, 2006).

There are three techniques within the replacement method that can be used depending on the research objective: namely, Permutation, Bootstrap, and Monte Carlo Simulation. Permutation uses sampling without replacement to test the hypothesis; Bootstrap uses sampling with replacement in finding confidence intervals; and Monte Carlo uses repeated sampling to the population whose properties have been identified for determining the sensitivity of statistical

procedures to the nature of the population (Berger, 2006). Even though each technique has its own purpose, the initial calculation procedure is the same. Before starting any analysis, other samples were produced using the original sample. This situation can be considered as a simulation that replaces the repetitive process of instrument implementation to the same population.

In finding the norm score distribution for this study, bootstrap method was chosen because the norm score of skills sought by this study was in confidence intervals. Furthermore, it is suitable to be used with a study that has a large sample size (Good, 2005; Berger, 2006). The basic technique of bootstrapping in getting confidence interval is the percentile interval. However, this technique was not used because there is other technique capable in estimating confidence interval with improved coverage (Statistic.com, 2009). The technique is known as the Bias-Corrected and Accelerated (BCA) bootstrap. The next section describes this technique in more detail.

4.3.3.1 BCA Bootstrap Technique

This technique has the same advantages such as the basic bootstrap in estimating confidence interval estimation to the desired parameters with the assumption that the original sample is able to represent the population. However, the significant differences between the two are with respect to implementation. In the basic bootstrap technique, random sampling with replacement is done to produce new samples that also have the same size as the original sample. Then, with the reproduction of many new samples, at least a thousand groups, the empirical sampling distribution can be produced, and thus confidence intervals can be set to

measure the desired statistics (Good; Berger, 2006). On the other hand, the BCA bootstrap technique is intended to improve the confidence intervals produced by the basic bootstrap. BCA bootstrap technique involves the use of the original sample and the basic bootstrap sample to estimate two other values i.e. the Bias and Acceleration values. Bias value consists of the difference between the true value of the population parameter being estimated and the median of the sampling distribution of that parameter, Acceleration value looks at "the degree to which the variance increases or decreases as the value of the population parameter increases" (Statistic.com, 2009). Both values are used in calculating a confidence interval.

Even though BCA Bootstrap is regarded as better than the basic bootstrap technique, its weakness is that it has a complex procedure that makes it hard to understand (Statistic.com, 2009). Stated below are the procedures that need to be employed in order to obtain the value of BCA Bootstrap, as recommended by Lunneborg (2000):

1. Compute the plug-in estimate, t, from the sample distribution, x.

2. Compute the n omitted-case estimates, $t_{(-i)}$.

3. Compute the average of the omitted-case estimates, $t_{(.)}$.

4. Compute the n influence statistics,

$$\frac{U_i}{n} = \left(\frac{n-1}{n}\right)\left(t_{(\cdot)} - t_{(-i)}\right)$$

5. Compute the acceleration estimate,

$$\hat{a} = \frac{\sum_{i=1}^{n}\left(\frac{U_i}{n}\right)^3}{6\left[\sum_{i=1}^{n}\left(\frac{U_i}{n}\right)^2\right]^{\frac{3}{2}}}$$

104

6. Form $\dot{X}$ from one or more copies of x.

7. Set $\hat{\theta}$ equal to the t of step 1.

8. Draw a bootstrap sample, x_b^*, from $\dot{X}$.

9. Compute the estimate, t_b^* , from the bootstrap sample.

10. Repeat steps 8 and 9 a total of B times, forming the bootstrap sampling distribution of t_b^*.

11. Compute $\hat{\pi}$, the proportion of the boostrap sampling distribution smaller than $\hat{\theta}$.

12. Select α, the confidence level for the $(1 - 2\alpha)100\%$ CI.

13. Use the tabled distribution of the standard normal distribution to determine $z_{[\theta]}$, $z_{[\alpha]}$, and $z_{[1-\alpha]}$. These are the z-scores that cut off the lower $\hat{\pi}$, α, and $(1 - \alpha)$ proportions of the standard normal distribution.

14. Use the acceleration estimate from step 5 and the z-scores from step 13 to compute

$$\hat{z}_{lo} = z_{[\hat{\pi}]} - \frac{\left(z_{[1-\alpha]} - z_{[\hat{\pi}]}\right)}{1 + \hat{\alpha}\left(z_{[1-\alpha]} - z_{[\hat{\pi}]}\right)}$$

and

$$\hat{z}_{up} = z_{[\hat{\pi}]} - \frac{z_{[\hat{\pi}]} - z_{[\alpha]}}{1 + \hat{\alpha}\left(z_{[\hat{\pi}]} - z_{[\alpha]}\right)}$$

15. Use the tabled distribution of the standard normal distribution to find q_{lo} and q_{up}, the proportions of the distribution falling below the two z-scores computed at step 14, $\hat{z}_{lo}$ and $\hat{z}_{up}$.

16. Use the two step 15 proportions and the B of step 10 to compute

$$lo = int\ [q_{lo} \times (B + 1)]$$

and

$$up = (B + 1) - int[(1 - q_{up}) \times (B + 1)]$$

17. Sort the bootstrap sampling distribution of step 10 in order from the smallest element, $t_{[1]}^*$, to the largest element, $t_{[B]}^*$.

18. Use the integers computed in step 16 to select $t_{[lo]}^*$ and $t_{[up]}^*$ from the sorted bootstrap sampling distribution. The selected $t_{[lo]}^*$ and $t_{[up]}^*$ are the q_{lo} and q_{up} quartiles of the bootstrap sampling distribution, just as $\hat{z}_{lo}$ and $\hat{z}_{up}$ were the q_{lo} and q_{up} quartiles of the standard normal distribution.

19. Steps 15 through 18 describe the translation of the CI limits from the standard normal to the bootstrap sampling distribution. The lower and upper limits of the $(1 - 2\alpha)100\%$ BCA confidence interval are given by $t_{[lo]}^*$ and $t_{[up]}^*$.

But with the existence of the Resampling Stats for Excel software to assist its calculation, the weakness of BCA Bootstrap has been overcome. A good reference on BCA Bootstrap can be found at Lunneborg (2000). Resampling Stats for Excel is an add-on software for Microsoft Office Excel produced by Resampling Stats, Inc. located in Arlington, Virginia, USA. It provides convenience in carrying out the simulation analysis of Bootstrap sampling and Permutation. It also allows analysis of even more than one million simulations and a sample iteration process performed at a time (Resampling Software, 2012). In this study, the mean scores of respondents for each skill were repeatedly measured ten thousand times. Subsequently, the confidence interval for each tested skill was obtained.

4.4 Summary

This chapter has discussed the six steps of instrument development involved to achieve the objectives of this study. The development of this instrument started with defining the instrument objective (Step one) and identifying the target population (Step two). These two steps are crucial to ensure that the developed instrument is valid. Next, the process involved selection of employability skills and measurement method to be used in the instrument. To achieve this, expert interviews were carried out to obtain the important elements that need to be measured on each selected skill. Items were then developed and presented to them for content validity. After that, the items were tested on a focus group for face validity. Then, the instrument was administered on UUM graduates in a pilot study. Based on the data of the pilot study, reliability test was run and improvement was made until the instrument was satisfactory. At this point, the first objective, which is to develop a valid and reliable instrument, was achieved. Next, before administering the instrument in the main study, the instrument was once again reviewed.

In the final step, the instrument was administered on final year students of UUM as the target population because of limitation in locating graduates. The scores obtained from the study were then extracted using EDA to get to know the data. Next, a hypothesis was tested on two independent variables to achieve the second objective, which is to ensure that the instrument is generic across quantitative and qualitative programmes of study. Finally, norm score distribution was developed using the resampling method of BCA Bootstrap as to achieve the third objective, which is to obtain norm scores that can be used to determine the level of employability skills of respondents. The next chapter discusses the result in further detail.

CHAPTER FIVE
RESULTS AND DISCUSSIONS

In this chapter, the discussion focus on achieving the objectives of (a) to develop an instrument to measure graduate employability skills of graduates, (b) to test the instrument across various fields of study, and (c) to obtain a norm score for the target population. The discussion starts from the validity and reliability of the instrument to the analyses made.

5.1 Instrument Validation

The following sections describe in detail the process of achieving validity and reliability of the instrument.

5.1.1 Validity of the instrument

In determining the validity of the instrument, content validity and face validity were considered. Expert opinions were solicited in determining content validity. They consist of four academicians of UUM. The detailed of those experts was included in Appendix B. Face validity was ascertained using a focus group consisting of individuals similar to the targeted respondents. The group was asked to answer the items in the instrument (initial draft) and give their opinions. Among the feedbacks received were the length of the questions and the format. Improvement were made where (a) the sentences were shortened without changing the purpose of the questions; and (b) the answering space was made closer to the item (was centralised in the pages next to the respected items).

As the item has been improvised, pilot test was conducted. A total of 500 copies of the instrument were distributed to graduates in the hall during the 23rd UUM convocation and they were asked to answer during the ceremony. These were collected at the end of the ceremony. In total, 334 (67%) graduates gave complete responses.

Table 5.1 showed the profile of respondents that involved with the pilot study. The respondents were between 22 and 52 years old. Of these, 289 respondents (86.5%) were 20s; and 67 respondents (24.1%) and 267 respondents (79.9%) were male and female, respectively. In terms of race, most of them were Malays (199 respondents or 59.6%) and Chinese (109 respondents or 32.6%). In terms of work experience, it ranged from one month to 32 years. Majority of the respondents seems to have work experience between one month to one year (183 respondents or 54.8%). With respect to programmes of study, there were 12 of these involved.

Prior to the scoring process, the data were checked for errors arising from not following instructions. As a result, no error was found in the pilot study and therefore, the data can be scored and analyzed.

One of the main aims of the pilot study is to seek feedback on the instrument. Of the total responses, 151 respondents gave their comments, with 114 of them acknowledging that the instrument was appropriate and 37 saying otherwise.

Table 5.1 Profile of Respondents in the Pilot Study

Age	20-29	289 (86.5%)
	30-39	27 (8.1%)
	40-49	11 (3.3%)
	50-59	7 (2.1%)
Gender	Male	67 (24.1%)
	Female	267 (79.9%)
Race	Malay	199 (59.6%)
	Indian	11 (3.3%)
	Chinese	109 (32.6%)
	Others	15 (4.5%)
Experience	No experience	91 (27.3%)
	≤ 1 year	183 (54.8%)
	$1 < x < 5$ years	34 (10.2%)
	$5 \leq x < 10$ years	12 (3.6%)
	≥ 10	14 (4.2%)

Of the latter, they commented that the instruction should be placed on the front page. Some mentioned that the purpose of the instrument should be expressed more clearly. Several indicated that the instrument was less effective because of the following reasons:

1. Questions and answers.

   a. Too long and need time to understand and answer questions.

   b. Answers should be limited to one choice only.

   c. Answering space of each item should be placed right after their response option.

   d. Questions were complicated and confusing.

   e. Response options were unclear.

   f. Instruction was too long.

   g. Less orderly.

2. Other comments/suggestions

   a. Not relevant to employability skills of graduates.

   b. Perfunctory answers.

   c. Scope of the questions narrow.

   d. Time to answer was inappropriate (timing was not correct)

Additionally, since there were respondents with considerable work experience, their comments were compared with that of those with low work experience to see whether there were differences. It appears that there is no relationship between work experience and the comments made as their respond were similar. Table 5.2 summarises the comments by work experience. There were negative and positive comments. Among the negative comments were, the instrument was claimed as complicated (the question are too long and confusing), the testing time was not suitable and the purpose of the instrument was not clear. As for the compliment, the instrument was said as an effective instrument to improve the quality of work, problem solving skill and to determine the quality of graduate.

Table 5.2 Opinions of Respondents by Experience

Work Experience (Years)	Respondents' Opinion
$x \geq 10$	1. Good instruments for positive thinking training.
	2. Good instruments, but the timing is not suitable because the respondents could not concentrate.
	3. The instruments are more suitable for young graduates.
	4. Questions and answers are too long.
	5. Instruments are boring.
$5 \leq x < 10$	1. Excellent as a guide to determine the quality of graduates.
	2. Can test the level of maturity in the face of the real world.
	3. Good instruments, but the questions are longer.
	4. More effective if the answering space is placed after the questions.
	5. There will be differences in responses if only one answer is to be selected.
	6. The study of graduate employability skills is not relevant.
$1 < x < 5$	1. Good and effective instruments.
	2. Very good to improve the quality of work.
	3. Testing the minds of graduates.
	4. Able to encourage graduates to think beyond the problem.
	5. Testing the level of rationality in decision making.
	6. Good as it gives some ideas for graduates to think of what they should do if the situation really happened.
	7. Instruments are quite good because there are no incorrect answers.
	8. Good instruments, but options are too long, and the answering method is confusing.
	9. Complicated instruments.

Table 5.2 Opinions of Respondents by Experience (Cont.)

Work Experience (Years)	Respondents' Opinion
$x \leq 1$	1. Good to expose graduates to the world of work.
	2. Attractive instrument.
	3. Instruments are good and very clear.
	4. Testing of the mind.
	5. Relevant and can serve as references when working.
	7. Able to assess the ability of graduates in solving problems.
	8. Help to get to know of yourself.
	9. Help produce more critical thinking.
	10. Good, but the format of response options should be changed to facilitate respondents to answer.
	11. Questions and response options are too long and require a long time to answer.
	12. Inappropriate time of distribution.
	13. More effective for graduates who have careers.
	14. Possibility for respondents to answer with care is low.
	15. The purpose of the instrument is less clear.

Modifications were then made on the instrument based on the comments. The original answering format was maintained to prevent respondents from making errors in choosing responses. Moreover, as explained in chapter three (Section 3.6.8), this format has the highest level of reliability compared with the other formats. In the main study, proper timing and location were considered and the instructions revised to make it clearer.

Next, the main study was conducted on second semester (2010/2011) to the final year UUM student as a substitute for graduate respondents because of the difficulty

of obtaining graduate due to once a year convocation ceremony. Out of 1200 copies of distributed instrument, 1012 (84%) were obtained. Based on Table 5.3, the respondents who completed the instrument were between 20 and 29 years old. Of these, 85.3% (863 respondents) of them were between 22 and 23 years. As for gender, most of them were females (816 respondents or 80.6%). For race, 579 respondents (57.2%) were Malay, 378 respondents (37.4%) were Chinese, and the rest of them were a small portion of Indian (33 respondents or 3.3%) and other races (22 respondents or 2.2%). As the copies of the instrument was distributed to the final year student (ranged from fifth to tenth semester), most responded instrument were from the sixth semester with 940 respondents (92.9%).

Table 5.3 Profile of Respondents in the Main Study

Age	20-23	863 (85.3%)
	24-26	127 (12.6%)
	27-29	22 (2.2%)
Gender	Male	196 (19.4%)
	Female	816 (80.6%)
Race	Malay	579 (57.2%)
	Indian	33 (3.3%)
	Chinese	378 (37.4%)
	Others	22 (2.2%)

Table 5.3 Profile of Respondents in the Main Study (Cont.)

Semester	5	42 (4.2%)
	6	940 (92.9%)
	7	5 (0.5%)
	8	23 (2.3%)
	9	1 (0.1%)
	10	1 (0.1%)
Field of study	Quantitative (10 programmes)	2894 (28.5%)
	Qualitative (21 programmes)	7267 (71.5%)

The data were checked for errors before further analysis. As a result, one error was detected in 42 copies of the instruments. The error showed that respondents have repeated answers in the same question (Example: A - B - C - A). In order to correct this error, telephone calls were made to the 42 respondents that had repeated answers for clarification. Following these interviews, three patterns of errors emerged. Each of the error patterns is best explained by using an example. In these examples, let us assume that the correct answer is "A – B – C – D":

1. An example of the first pattern (Figure 5.1) was an answer in the form of "A - B - C - A", where "A" was repeated for extreme answers. Six respondents made such an error due to their carelessness when answering the item.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
Contoh Soalan	A	B	C	A	

Figure 5.1 First Pattern Error

2. The second pattern (Figure 5.2) shows the difference where the repetition was made. An example of this pattern is "A - B - A - C", where "A" was first placed as being the Most Effective Action and the second "A" was placed as being an Ineffective Action. In this situation, the respondents felt that the most effective action can also be perceived as an ineffective action depending on the circumstances. Ten respondents made this error.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
Contoh Soalan	A	B	A	C	

Figure 5.2 Second Pattern Error

3. The last pattern of answer arrangement is "A - A - C - D" (Figure 5.3). Compared to the first two patterns, this pattern was the most error, committed by the respondents (36 respondents). The respondents made this error because they think that among the four response options provided, three of them were ineffective actions. Actually, as mentioned earlier, this study provided four columns that consist of two columns for effective action and another two columns for an ineffective action (refer Table 4.3). So, there's no room for the respondents to place his third ineffective action in the answering space. Hence, in answering the respective item, they repeatedly place the

same answer of effective action in the two spaces provided for an effective action.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
Contoh Soalan	A	A	C	D	

Figure 5.3 Third Pattern Error

Based on the interview, this study found that the errors were because the respondents did not understand the instruction provided in the instrument. They were not told that when arranging the answers, every response option must be written down and that they could write each letter once. Thus, to solve the problem, the instruction was modified by clearly stating what respondents need to do. The error contained in the 42 sets of instrument were fixed throughout the interview. Once the data had been confirmed to be free of error, the scoring process was implemented.

5.1.2 Reliability of the Instrument

As the scores of pilot and main studies had been obtained, this study proceeded with reliability analysis. Showed below are Figure 5.4 to Figure 5.7 that comparing boxplot, Exploratory Factor Analysis (EFA) and Cronbach's Alpha value for each tested skill in this study.

As it can be seen Figure 5.4, the location of boxplot for Item One in the pilot and main study are higher than Item Seven. This indicate that respondents score higher mark for Item One as compared to Item Seven for Communication skill. This also indicates that both items may have a different level of difficulty. Item Seven seems

117

to be more difficult than Item One. This perhaps contribute to low factor loading in EFA and Cronbach's Alpha and shows that those items are not reliable based on internal consistency. However, if they are compared item by item (i.e.: Item One with Item Seven) across studies (pilot and main), both items shows a similar score.
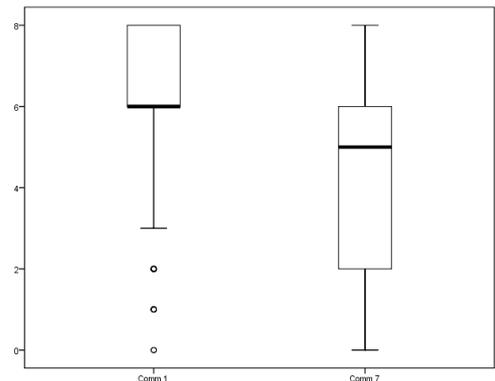
Item	Scores distribution	EFA	Cronbach's α
Pilot Study — One		0.033	
			0.064
Pilot Study — Seven		0.254	
Main Study — One		0.181	
			0.028
Main Study — Seven		0.192	
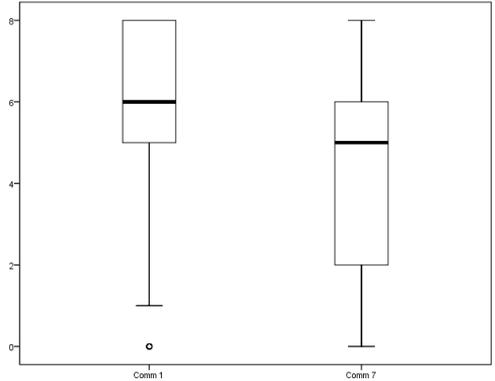
Figure 5.4 Comparison of Boxplot, EFA and Cronbach's Alpha for Items of Communication Skill

Figure 5.5 shows boxplot, EFA and Cronbach's Alpha of Item Two and Ten, which was used to measure Professional work ethics and morality. Based on the figure, the score distribution of Item Ten for the pilot and main study is wider compared to Item Two in both studies. For Item Ten, most of the respondents obtained scores ranged from four to seven marks. As four Item Two, there are two conditions. For the pilot study, the score of most respondents ranged from four to six marks. As for the

main study, only five and six marks shows a high frequency of score. This indicates that the difficulty level of Item Ten was lower compared to Item Two as in both studies (pilot and main), many respondents able to get seven marks. Perhaps, this factor may contribute to low factor loading of EFA and Cronbach's Alpha. As for the comparison on item by item across studies, both items shows consistent scores but with a slight differences of scores in Item Two.

	Item	Scores distribution	EFA	Cronbach's α
Pilot Study	Two		-0.628	
				0.26
	Ten		-0.194	
Main Study	Two		0.577	
				0.132
	Ten		0.075	

Figure 5.5 Comparison of Boxplot, EFA and Cronbach's Alpha for Items of Professional Work Ethics and Morality

Next, Figure 5.6 shows the scores of four items measuring Entrepreneurial skill. It can be seen that the location of boxplot for each item were different, both in pilot study and main study. The boxplot for Item Three is located at the lowest score. As for Item Four and Eight, the boxplot for both items are located at the middle of the

119

graph indicating average scores was mostly obtained. Finally, for Item Eleven, the location of boxplot is the highest compared to other items. This indicates that the easiest item to be scored is Item Eleven and the hardest item to be scored is item six. As for Item Four and Eight, their difficulty level are on average. As usual, perhaps these differences may contribute to low factor loading of EFA and Cronbach's Alpha. However, across studies, comparison of item by item shows a consistent scoring pattern for each item.

	Item	Scores distribution	EFA	Cronbach's α	α if item deleted
Pilot Study	Three		0.668		-0.017
	Four		-0.101		0.177
				0.048	
	Eight		0.044		-0.040
	Eleven		0.248		0.029
Main Study	Three		0.168		0.021
	Four		-0.310		0.108
				0.018	
	Eight		0.058		-0.100
	Eleven		0.337		0.024

Figure 5.6 Comparison of Boxplot, EFA and Cronbach's Alpha for Items of Entrepreneurial Skill

Subsequently, as Figure 5.6, the patterns of boxplot for each item that measures Personal quality in Figure 5.7 shows variability in their location and this may affect the value of factor loading in EFA and Cronbach Alpha to become low. For both studies of pilot and main, boxplot with the lowest location is on Item Three. Conversely, as for Item Five and Eleven, both item shows the highest location of boxplot in both studies. As for Item Two, Four, Eight and Ten, each of them display an average position of boxplot.

Item	Scores distribution	EFA	Cronbach's α	α if item deleted
Pilot Study				
Two		0.091		0.109
Three		0.668		0.206
Four		-0.101		0.239
Five		-0.005	0.2	0.222
Eight		0.044		0.166
Ten		-0.135		0.123
Eleven		0.248		0.158
Main Study				
Two		0.557		0.109
Three		0.157		0.157
Four		0.099		0.181
Five		-0.034	0.150	0.129
Eight		0.022		0.098
Ten		0.075		0.111
Eleven		0.033		0.130

Figure 5.7 Comparison of Boxplot, EFA and Cronbach's Alpha for Items of Personal Quality

121

In term of difficulty, it can be said that may be the most difficult item in Personal quality is Item Three and the easiest item to be scored comprises of Item Five and Eleven. Although the measurement of this skill involves a lot of item with differences in their difficulty level, the distribution of scores by comparing item by item are consistent across studies.

Finally, Figure 5.8, which consist of items that was used to measure Critical thinking in problem solving, also display items with different location of boxplot.

Item	Scores distribution	EFA	Cronbach's α	α if item deleted
Pilot Study				
Six		-0.261		0.021
Nine		-0.013	-0.092	-0.111
Twelve		0.309		-0.071
Main Study				
Six		0.489		-0.007
Nine		-0.027	-0.010	0.016
Twelve		-0.125		-0.036

Figure 5.8 Comparison of Scores Distribution, EFA and Reliability for Items of Critical Thinking in Problem Solving

Item Six shows that the location of its boxplot was higher compared to Item Nine but equal to Item Twelve. This indicates that many respondent able to get high marks in Item Six and Twelve and inversely moderate marks in Item Nine. Consequently, it may causing the factor loading of EFA and Cronbach's Alpha to become low. Even so, the scores of item by item comparison across studies for each item is consistent.

Next, Table 5.4 displayed the percentage of scores for the pilot and main studies according to employability skills. The frequency of scores were converted to percentage because there was a big differences in the sample size for both studies (334 and 1012 respondents, respectively), the conversion will allow the scores for both studies to be compared. As displayed in Table 5.4, the bolded and italic numbers are the highest percentages that represent the most scored marks by respondents in each item. Those percentages were consistent with the location of the boxplot showed in Figure 5.4 until Figure 5.8. For Communication skill, the score percentage for Item Seven is lower than Item One indicating that Item Seven has a higher difficulty level compared to Item One. For Professional ethics and morality, Item Two shows lower percentage compared to Item Ten. So, Item Ten can be said as easier to be scored compared to Item Two. As for Entrepreneurial skill, Item Three shows the lowest percentage of score, Item Eleven with the highest and Item Four and Eight were in the middle. Therefore, for this skill, Item Three can be said as more difficult than the other three items. While for Item Eleven, its difficulty level seems to be low. Next, for Personal quality, it have three items with low difficulty (Item Five, Ten and Eleven), three item with an average difficulty (Item Two, Four and Eight) and one fairly difficult item, namely, Item Five. Finally, for

Table 5.4 Percentages of Scores for Pilot and Main Study

Skill	Item	Study	Score								
			0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Communication Skill	One	Pilot	0.6	2.1	6.0	0.9	9.9	4.5	29.6	3.9	*42.5*
		Main	1.2	2.6	5.4	1.7	10.8	3.9	31.6	3.6	*39.3*
	Seven	Pilot	11.4	6.3	10.5	4.5	9.0	18.0	*21.9*	9.6	9.0
		Main	13.3	7.9	8.1	4.8	7.5	19.0	*21.7*	9.5	8.1
Professional Work Ethic and Morality	Two	Pilot	0.9	2.1	15.0	2.1	5.4	6.6	*48.8*	1.5	17.7
		Main	0.4	2.1	12.8	1.7	4.8	6.6	*50.6*	3.0	18.1
	Ten	Pilot	2.1	7.2	7.8	6.3	6.9	11.1	18.6	15.6	*24.6*
		Main	2.7	5.9	9.1	6.7	7.8	9.3	19.5	17.2	*21.8*
Entrepreneurial Skill	Three	Pilot	21.6	10.2	*26.7*	13.2	21.9	0.9	2.7	1.2	1.8
		Main	20.3	9.2	*25.4*	15.9	23.2	0.4	2.5	1.0	2.2
	Four	Pilot	12.0	9.9	16.2	4.8	9.9	7.8	*31.1*	3.3	5.1
		Main	14.6	9.4	15.7	4.7	8.9	8.1	*27.8*	3.3	7.5
	Eight	Pilot	2.1	3.0	16.5	3.0	12.6	3.0	*46.4*	0.9	12.6
		Main	2.5	3.0	17.0	3.5	12.8	2.8	*43.6*	0.8	14.2
	Eleven	Pilot	0.6	3.3	3.3	2.1	14.4	10.2	20.1	9.0	*37.1*
		Main	0.2	3.7	3.9	2.8	13.6	10.6	21.4	7.2	*36.7*
Personal Quality	Two	Pilot	0.9	2.1	15.0	2.1	5.4	6.6	*48.8*	1.5	17.7
		Main	0.4	2.1	12.8	1.7	4.8	6.6	*50.6*	3.0	18.1
	Three	Pilot	21.6	10.2	*26.7*	13.2	21.9	0.9	2.7	1.2	1.8
		Main	20.3	9.2	*25.4*	15.9	23.2	0.4	2.5	1.0	2.2
	Four	Pilot	12.0	9.9	16.2	4.8	9.9	7.8	*31.1*	3.3	5.1
		Main	14.6	9.4	15.7	4.7	8.9	8.1	*27.8*	3.3	7.5
	Five	Pilot	0.6	1.2	8.1	8.4	20.1	2.1	3.0	8.7	*47.9*
		Main	0.7	2.0	8.2	8.5	22.4	1.6	2.7	8.7	*45.3*
	Eight	Pilot	2.1	3.0	16.5	3.0	12.6	3.0	*46.4*	0.9	12.6
		Main	2.5	3.0	17.0	3.5	12.8	2.8	*43.6*	0.8	14.2
	Ten	Pilot	2.1	7.2	7.8	6.3	6.9	11.1	18.6	15.6	*24.6*
		Main	2.7	5.9	9.1	6.7	7.8	9.3	19.5	17.2	*21.8*
	Eleven	Pilot	0.6	3.3	3.3	2.1	14.4	10.2	20.1	9.0	*37.1*
		Main	0.2	3.7	3.9	2.8	13.6	10.6	21.4	7.2	*36.7*
Critical Thinking in Problem Solving	Six	Pilot	0.6	1.8	12.0	3.0	25.8	1.8	14.7	3.3	*37.1*
		Main	0.9	2.8	12.2	2.5	26.7	2.8	12.9	5.5	*33.8*
	Nine	Pilot	3.0	2.4	30.8	4.8	*50.0*	5.4	0.6	3.0	3.0
		Main	2.8	2.2	34.6	4.7	*47.9*	0.2	4.3	0.5	2.9
	Twelve	Pilot	3.9	1.2	4.8	3.0	7.5	4.5	5.7	*36.2*	33.2
		Main	4.2	1.7	3.3	3.0	6.7	4.5	5.7	*37.9*	33.1

the items of Critical thinking in problem solving, Item Six and Twelve shows low difficulty, while Item Nine shows a moderate level of difficulty. These scores seems to be consistent with the tabulation of boxplot in Figure 5.4 to Figure 5.8.

As for the scores across studies, the percentage of scores between the pilot and the main study seems to be similar based on item by item across the pilot and the main study. To confirm this similarity, Chi-Square test for homogeneity was conducted. Shown below are the hypothesis use in conducting this test:

$H_{0:}$ The percentage of scores for each item in the pilot and the main study are homogenous

$H_{1:}$ The percentage of scores for each item in the pilot and the main study are not homogenous

Table 5.5 indicated that all of the p-values were not significant. Therefore the percentage of scores of each item in the pilot and the main study was homogeneous. This perhaps explained the low factor loading and Cronbach's Alpha for each skills and jeopardised the reliability (internal consistency) of items in the construct (skills). Although the items are not reliable, but they were consistent stable based on item by item. Additionally, based on the content validity of the experts, all of the items developed in this study are measuring their respected skill. Therefore, each item will remained in the instrument. As a matter affect, items with different level of difficulties can be used to differentiate average respondent and excellent respondent according to the tested skills. This study did not achieve the internal consistency due to level of difficulty among items in the same skill but the grouping of the items

125

were valid based on content validity by the experts. Therefore the data from the instrument are suitable for further analysis.

Table 5.5 Chi-Square Test for Homogeneity

Employability Skills	Item	Asymp. Sig. (2-sided)
Communication skill	One	0.979
	Seven	0.967
Professional ethics and morality	Two	0.953
	Ten	0.993
Entrepreneurial skill	Three	0.973
	Four	0.991
	Eight	0.992
	Eleven	0.992
Personal quality	Two	0.953
	Three	0.973
	Four	0.991
	Five	0.962
	Eight	0.992
	Ten	0.993
	Eleven	0.992
Critical thinking in problem solving	Six	0.968
	Nine	0.468
	Twelve	0.999

Throughout this section, this study has able to achieve the first objective of the study, which was to obtain an instrument that is valid (was proven through content and face validity) and reliable in term of consistency stability (was proven through the comparison of boxplot, percentage of scores and Chi-Square test). Next section discuss about the process of achieving the second objective, which was to obtain a generic instrument across graduates.

5.2 Items' Practicability Across Different Fields of Study

This section was intended to achieve the second objective, which was to develop a generic employability skills instrument for graduates across undergraduate programmes. Mann-Whitney U test was run to test the existence of differences in the scores between two fields of study (ACNielsen, 2000):

$H_{0:}$ There are no differences in scores between respondents in quantitative and qualitative fields of study.

$H_{1:}$ There are differences in scores between respondents in quantitative and qualitative fields of study.

Table 5.6 shows no differences in the scores between the respondents of quantitative and qualitative fields of study ($p < .05$). Hence, based on these results, it is appropriate to conclude that the instrument in this study is generic across quantitative and qualitative fields of study.

Next section discusses the analysis conducted to achieve the third objective of this study, which is to obtain a norm for the target population.

Table 5.6 Results of Mann-Whitney Test on Score Differences across Field of Study

Employability Skills	Item	Asymp. Sig. (2-tailed)
Communication skill	One	0.886
	Seven	0.223
Professional ethics and morality	Two	0.491
	Ten	0.060
Entrepreneurial skill	Three	0.061
	Four	0.868
	Eight	0.337
	Eleven	0.532
Personal quality	Two	0.491
	Three	0.061
	Four	0.868
	Five	0.998
	Eight	0.337
	Ten	0.060
	Eleven	0.532
Critical thinking in problem solving	Six	0.053
	Nine	0.328
	Twelve	0.051

5.3 Norm Score Development

Before calculating the norm score, this study had first developed the scoring. Table 5.7 shows the level of each employability skill possessed by the respondents. Overall, more than 50% of the respondents possessed a moderate to high level of employability skills except for the Entrepreneurial Skills, whereby only 38.1% (386) possessed a moderate to high level. Majority (69.3%) possessed a moderate to high level of Professional Ethics and Morality skill, and 69.1% (699) had Critical Thinking in Problem Solving. By using these scores, norm scores can be calculated.

As mentioned earlier in Section 4.3.3.1 in Chapter Four, the developed norm score consists of the value of confidence interval for each tested skill. BCA bootstrap was used to calculate the norm and to assist the calculation process. Resampling Stats for Excel software was used to conduct a simulation on the mean score of the five employability skills based on the score obtained in the main study. The simulation was run to 10000 repetitions using 95% confidence level. As a result, this study has been successfully obtained the norm score for each tested skill.

Table 5.8 shows the summary of the norm score obtained from the resampling procedure. For Communication skill, the norm score was between 5.08 and 5.28. The norm score for Professional ethics and morality was between 5.34 and 5.54. The norm score for Entrepreneurial skill was between 4.22 and 4.35. For Critical thinking in problem solving, the norm score was between 4.94 and 5.08. The norm score of Personal quality was between 4.78 and 4.89.

Table 5.7 Employability Skills Level Possessed by the Respondents

Employability Skills	Level	Frequency	Valid Percent
Communication skill	Low	395	39.0
	Moderate	423	41.8
	High	194	19.2
	Total	1012	100.0
Professional ethics and morality	Low	311	30.7
	Moderate	463	45.8
	High	238	23.5
	Total	1012	100.0
Entrepreneurial skill	Low	626	61.8
	Moderate	382	37.7
	High	4	0.4
	Total	1012	100.0
Personal quality	Low	350	34.6
	Moderate	659	65.1
	High	3	.3
	Total	1012	100.0
Critical thinking in problem solving	Low	313	31.0
	Moderate	674	66.6
	High	25	2.5
	Total	1012	100.0

Low: 0.00 to 2.66; Moderate: 2.67 to 5.33; High: 5.34 to 8

Based on the score guideline in Table 4.7 in Chapter Four, on average, the final year local undergraduate students of Universiti Utara Malaysia (UUM) had a moderate level of Communication skill, Entrepreneurial skill, Critical thinking in problem solving, and Personal quality. As for Professional ethics and morality, the norm scores indicated that the average performance of the respondents for this skill was high. Even though the norm scores developed can be used as a benchmark in determining the level of skills possessed by UUM graduates in the future, they need to be occasionally updated and revised to reflect changes in skills that may ensue.

Table 5.8 Norm Score for Graduate Employability Skills Instrument

Employability Skills	BCA 95% LCL	BCA 95% UCL
Communication skill	5.08	5.28
Professional ethics and morality	5.34	5.54
Entrepreneurial skill	4.22	4.35
Personal quality	4.78	4.89
Critical thinking in problem solving	4.94	5.08

Low: 0.00 to 2.66; Moderate: 2.67 to 5.33; High: 5.34 to 8

5.6 Summary

In summary, all the objectives proposed in this study have been successfully achieved. Firstly, by taking into consideration expert opinion, focus group, and implementation of Chi-Square test for homogeneity, a valid and reliable employability skills instrument has been successfully developed. The instrument has also shown to be generic across graduates as the scores of graduates in different fields of study (Qualitative and Quantitative) demonstrated no difference. Finally,

this study has also successfully obtained norm scores for each employability skill. By using these scores, a better interpretation on the individual score on the respective skill can be made.

CHAPTER SIX
CONCLUSION AND RECOMMENDATIONS

This study sought to achieve three objectives: (1) to develop an instrument to measure graduate employability skills of graduates.; (2) to test the instrument across various fields of study; and (3) to obtain a norm score for the target population.

Specifically, each objective was successfully achieved. The instrument was developed with 12 items that consist of five skills i.e. communication skill, personal quality, critical thinking in problem solving, professional ethics and morality, and entrepreneurial skill. The situational judgment test (SJT) was used to measure the skills. The validity of the instrument was ascertained through expert opinions (Content validity) and focus group (Face validity). Based on the feedback given, modifications to the instrument were made. Chi-Square test for homogeneity was used to determine the reliability by testing the homogeneity of score for each tested item between the pilot and the main study.

As for the second objective, hypothesis testing was conducted to see the differences in scores between the quantitative and qualitative programmes. As the results showed no significant differences and may support the argument that the instrument is generic.

Finally, to achieve the third objective, norm score distribution was developed using the resampling method of BCA Bootstrap. By using this method, an interval of norm

score for each tested skill was generated to be used to describe the level of employability skill of respondents.

6.1 Research Implication

As mentioned earlier, the instrument was administered to UUM final year students. Results indicate that 50% of them had at least a moderate level of each employability skill. However, only 30% of them had moderate and good Entrepreneurial skills. The result identifies that the current final year students of UUM may be weak in their Entrepreneurial skills, and measures need to take to improve this skill level. For example, the university could think of ways on how to enhance this type of skill amongst its graduates so that they are employable in the future. Besides graduates and the university, this instrument can also benefit employers, government agencies and career development agencies:

1. Employers.

   Replace an initial interview stage and speed up the screening process in finding potential employees.

2. Government agencies

   Reduce the cost and time required in graduate training program without affecting the quality of training.

3. Career development agency

   Assist in placing graduates in an appropriate organization based on the skills they possess.

6.2 Limitation

One of the important considerations that need to be taken into account in developing an instrument is its validity and reliability. While the validity of the instrument was successfully ascertained through expert opinion, respondents' view, and the meta-analysis of previous studies on the validity of the SJT, determining the reliability was more a challenging despite the improvements made on the instrument. But such situation is normal for the SJT instrument, as indicated by previous research especially due to the multidimensional nature of SJT. Additional, it has been found that it is important to set up a consistent difficulty level among item as it may also influence the reliability of the instrument. This criteria has not being considered in developing the item in this study. So, it is expected that this is another main things that affect the reliability of the instrument in this study. Despite this setback, this study employed the Chi-Square test for homogeneity to check for homogeneity in the scores across the pilot and the main study. As a result, each item shows that their scores are consistent.

The second limitation was that the instrument was not administered to the actual graduates; rather final year students of UUM were used. Graduates were considered in the pilot study only. Because graduates were located across the country, and because of other limitations that constrained the main study, final year students of UUM were considered as the target respondents. Even though final year students and graduates were not identical, the former was the closest representation of the latter, as they were more or less similar to those who have graduated. In addition, based on the evaluation made by the respondents, the original level of employability skills appeared not to be influenced by the working environment. If the respondents were

graduates, there is a possibility that the projected level of employability skills would be influenced by work experience at least during the industrial training.

However, despite these limitations, there is still room for improvement on the instrument to make sure that is ready to be administered on a larger scale.

6.3 Suggestions for Future Research

Based on the limitations above, a few suggestions can be offered to future researchers, as follows, 1) considering the difficulty level of items in developing SJT instrument, 2) develop a statistical method to determine reliability that is suitable for SJT, 3) develop more SJT items for the employability skills to create an item inventory, 4) further improve the instrument and administer it at other universities in an effort to test the generic capability of this instrument across graduates in Malaysia, 5) develop an online instrument and 6) periodically review the norm score distribution.

REFERENCES

About MCQBank (2012). Retrieved January 17, 2012, from http://mcqbank.co.uk/about-us.

ACNielsen Research Service (2000). *Employer satisfaction with graduate skills*. Retrieved on 14 Mei 2009 from www.dest.gov.au/archive /highered/eippubs/eip99-7/eip99_7pdf.pdf

About Resampling (2012). Retrieved January 31, 2012, from http://www.resample.com

ACT (2008). *WorkKeys: Preparing for the WorkKeys Assessments*. Retrieved from www.pennfield.net/hs/docs/workkeys.pdf

Adult Education (2011). Retrieved December 26, 2011, from http://www.southark.edu/index.php/continuing-education/adulteducation

Alicea, N.R. & Scott, G. (2007). *A survey of selected work readiness certificates. Job For the Future*. Retrieve on 23 Jun 2009 from http://www.jobsfirstnyc.org/docs/WorkReadiness.pdf

Armstrong, G. (2005). Managing performance through people: The challenge for tomorrow's organization. In Cooper, C. L. (Ed.), *Leadership and management in the 21st century: Business challenges of the future* (pp. 285). New York: Oxford University Press.

Berger, D. (2006, April). *Introduction to resampling techniques*. Statistical Workshop presented at the Western Psychological Association Palm Springs, California. Retrieved 25 April 2011 from http://wise .cgu.edu/downloads/Introduction%20to%20Resampling%20Techniques%2006 0420.doc

Bergman, M. E., Drasgow, F., Donovan, M. A., & Henning, J. B. (2006). Scoring situational judgment tests: Once you get the data, your troubles begin. *International Journal of Selection and Assessment, 14*(3), 223-235. doi: 10.1111/j.1468-2389.2006.00345.x

Bess, T. L. (2001). *Exploring the dimensionality of situational judgment: Task and contextual knowledge*. Unpublished master's dissertation, Faculty of Virginia Polytechnic and State University.

Bizeur, J. L., Farchy, I., & De Maen, P. (2010). *Situational judgement test for european institution competitions: A methodology and 60 questions with answers*. Brussels: Orseu.

Boatwright, J. R., & Slate, J. R. (2002). Development of an instrument to assess work ethics. *Journal of Industrial Teacher Education, 39*(4). Retrieved 8 January 2011 from http://scholar.lib.vt.edu/ejournals/JITE/v39n4/boatwright.html

Breaugh, J., & Starke, M. (2000). Research on employee recruitment: so many studies, so many remaning questions. *Journal of Management, 26*(3), 405-434. doi: 10.1177/014920630002600303

Bridgstock, R. (2009). The graduate attributes we've overlooked: Enhancing graduate employability through career management skills. *Higher Education Research & Development, 28*(1), 31-44. doi: 10.1080/07294360802444347

Bridgstock, R. (2011). Skills for creative industries graduate success. *Education + Training, 53*(1), 9 – 26. doi: 10.1108/00400911111102333

Cardwell, M. (2003). *A-Z psychology*. New York: McGraw-Hill.

Career Center. (2006). *Identify your personal qualities*. Retrieved on 8 January 2011 from http://www.sc.edu/career/Pdf/identifypersonalqualities.pdf

Careers in South East Asia / Recruitment Process (2011). Retrieved January 15, 2012, from http://www.adlittle.com.my/recruitment-process_my.html

CASAS (20112). Retrieved January 15, 2012, from https://www.casas.org/about-casas

Coakes, S. J., & Steed, L. (2007). *SPSS: analysis without anguish - Version 14.0 for Windows*. Milton: John Wiley & Sons.

Cohen, R. J., & Swerdlik, S. E. (2009). *Psychological testing and assessment: An introduction to tests and measurement* (7th ed.). Columbus: McGraw-Hill

Colton, D., & Covert, R. W. (2007). *Designing and constructing instruments for social research and evaluation*. San Francisco: John Wiley & Sons, Inc.

Commonwealth of Australia (2002). *Employability skills for the future.* Retrieved on 25 April 2009 from http://www.dest.gov.au/sectors/training_skills /publications_resources/pr.ofiles/documents/final_report_x1_pdf.htm

Cooper, C. L. (Ed.). (2005). *Leadership and management in the 21st century: Business challenges of the future*. Oxford: Oxford University Press.

Chan, D., & Schmitt, N. (2002), Situational judgment and job performance. *Human Performance, 15*(3), 233-254. doi: 10.1207/s15327043hup1503_01

Craighead, W. E., & Nemeroff, C. B. (2004). *The concise Corsini encyclopedia of psychology and behavioral science* (3rd ed.). New Jersey: John Wiley & Sons.

Chisholm, D., Knapp, M. R. J., Knudsen, H. C., Amaddeo, F., Gaite, L., Bob Van Wijngaarden, B. V., & The Epsilon Study Group (2002). Client socio-demographic and service receipt inventory - European version: Development of an instrument for international research: EPSILON Study 5. *The British Journal of Psychiatry,* BJP 2000, 28-33 doi: 10.1192/bjp.177.39.s28

Curtis, D. & McKenzie, P. (2001). Employability skills for Australian industry: Literature review and framework development. *Employability skills for future project 2002*. Retrieved on 14 Mei 2009 from http://www.dest.gov.au/archive/ty/publications/employability_skills/literature _research.pdf

Dennis, R. S., & Bocarnea, M. (2005). Development of the servant leadership assessment instrument. *Leadership & Organization Development, 26*(8), 600-615. doi: 10.1108/01437730510633692

Deraman, N., Mohamad, A., Bakar, H., Hashim, N., & Keat, O. Y. (2010). *Keusahawanan: Teori dan praktis* (2nd ed.). Kuala Lumpur: McGraw-Hill.

Department of Statistics (2008). *Laporan kaji selidik tenaga buruh suku tahun ke-4 2007* Retrieved on 6 Mac 2009 from http://www.statistics.gov.my/ portal/download_Labour/download.php?file=BPTMS/report_ptb_q42007.pdf.

Department of Statistics (2011). *Laporan kaji selidik tenaga buruh suku tahun ke-4 2010.* Retrieved on 8 Mei 2011 from http://www.statistics.gov.my/ portal/download_Labour/download.php?file=BPTMS/report_ptb_q42010.pdf.

DeVellis, R. (2003). *Scale development: Theory and applications*. London: Sage

Devito, J. A. (2009). *Human communication: The basic course.* New York: Allyn & Bacon.

Dunning, D., Meyerowitz, J. A., & Holzberg, J. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, *57*(6), 1082-1090.

Fallon, J. D., Avis, J. M., Kudisch, J. D., Gornet, T. P., & Frost, A. (2000). Conscientiousness as a predictor of productive and counterproductive behaviors. *Journal of Business and Psychology, 15*(2) 339-349.

FAQs (2011). Retrieved December 26, 2011, from http://www.workreadiness.com/faqs.html

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage Publications.

Fielding, N. G. (2006). Self-Report study. In V. Jupp (Ed.). *The Sage dictionary of social research methods (vol. 1&2)* (pp. 275-277). Great Britain: Sage Publications.

Filliben, J. J. (2003). Exploratory data analysis. In C. Croarkin & P. Tobias (Eds.), *NIST/SEMATECH e-handbook of statistical methods.* Retrieved on 6 March 2011 from http://www.itl.nist.gov/div898/handbook/eda/eda.htm

Garman, A. N., Johnson, M. P., & Howard, D. M. (2006, May). *Development of a situational judgment test to assess educational outcome.* Poster presented at the annual meeting of the Society of Industrial / Organizational Psychologists, Dallas, Texas. Retrieve on 3 January from http://liacus.com/free_resources/HALA_siop2006.pdf.

Geren, B. (2011). The work ethic: Is it universal? *Journal of International Business and Cultural Studies, 5.* 1-9. Retrieved on 8 January 2012 from www.aabri.com/manuscripts/10678.pdf

Good, P. I. (2005). *Introduction to statistics through resampling methods and Microsoft Office Excel.* New Jersey: John Wiley & Sons.

Grubb, L., III. (2005). *The emotional side of a situational judgment test.* Retrieved on 30 May 2010 from http://www.westga.edu/~bquest/2005/ emotional.pdf

Guion, R. M., & Highhouse, S. (2006). *Essentials of personnel assessment and selection.* New Jersey: Lawrence Erlbaum Associates.

Gupta, M., Boyd, L., & Kuzmits, F. (2011). The evaporating cloud: a tool for resolving workplace conflict. *International Journal of Conflict Management. 22* (4). 394-412. doi: 10.1108/10444061111171387

Halpern, D. F. (2003). *Thought and knowledge: An introduction to critical thinking* (4th ed.). New Jersey: Lawrence Erlbaum Associates.

Hanson, M. A., Borman, W. C., Mogilka, H. J., & Manning, C. (1999). Computerized assessment of skills for a highly technical job. In F. Drasgow, & J. B. Olson-Buchanan (Eds.), *Innovations in computerized assessment* (pp. 197-220). New Jersey: Lawrence ErlBaum Associates.

Harris, A. (2011). *Deconstructing Workplace Conflict Resolution.* Unpublished master's dissertation, Auckland University of Technology.

Hashim, N. (2009). *The public higher learning institution students' and its internal environment readiness towards entrepreneurship education.* Unpublished doctoral thesis, Universiti Kebangsaan Malaysia.

Huang, F., & Cappelli, P. (2006). Employee screening: Theory and evidence. *SMU Economics & Statistics Working Paper Series.* Retrieved on 4 January 2011 from http://www.nber.org/papers/w12071.

Hodges, D., & Burchell, N. (2003). Business graduate competencies: Employers' views on importance and performance. *Asia Pacific Journal of Cooperative Education, 4*(2), 16-22.

Hogan, J., & Ones, D. S. (1997). *Conscientiousness and integrity at work.* In R. Hogan, J. A. Johnson, & S. R. Briggs, (Eds.). *Handbook of personality psychology* (pp. 849-870). San Diego: Academic Press. doi: 10.1016/B978-012134645-4/50033-0

Holbrook, A. (2008). Self-Reported measure. In P. J. Lavrakaz (Ed.), *Encyclopedia of survey research methods* (pp. 804-805). United States of America: Sage Publications.

Hooper, A. C., Cullen, M. J., & Sackett, P. R. (2006). *Operational threats to the use of SJTs: Faking, coaching and retesting issues.* In J. A. Weekley, & R. E. Polyhart, R. E. (Eds.), *Situational judgment tests: theory, measurement and application* (pp. 205-232). New Jersey: Lawrence Erlbaum Associates, Inc.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport: Praeger.

Jesiek, B. K., & Woo, S. E. (2011). *Realistic assessment for realistic instruction: Situational assessment strategies for engineering education and practice.* Proceedings from WEE2011 (pp. 205-212). Retrieved on 29 December 2011 from http://www.sefi.be/wp-content/papers2011/T7/108.pdf

Joiner, D. A. (2008, October). *Situational judgment tests.* Paper presented at Western Region Intergovernmental Personnel Assessment Council, Sparks, Nevada. Retrieved on 3 May 2010 from http://www.wripac.org/files/SituationalJudgmentTests.pdf

Kalton, G. (1983). *Introduction to survey sampling.* California: Sage Publication.

Kane, M., Berryman, S., Goslin, D., & Meltzer, A. (1990). Identifying and describing the skills required by work. Retrieved on 18 Feb 2009 from http://wdr.doleta.gov/SCANS/idsrw/idsrw.pdf

Kisamore, J. L., Jawahar, I. M., Liguori, E. W., Mharapara, T. L., & Stone, T. H. (2010). Conflict and abusive workplace behaviors: The moderating effects of social competencies. *Career Development International, 15* (6), 583-600. doi: 10.1108/13620431011084420

Legree, P. J., & Psotka, J. (2006). *Refining situational judgment test methods. Retrieved on 5 January 2011 from http*://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA481655.

Leighton, J. P., & Gierl, M. J. (2011). *The learning sciences in educational assessment: The role of cognitive models.* New York: Cambridge University Press

Lievens, F., & Coetsier, P. (2002). Situational tests in student selection: An examination of predictive validity, adverse impact, and construct validity. *International Journal of Selection and Assessment, 10*(4), 245-257.

Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment test: A review of recent research. *Personnel Review*, *37*(4), 426 - 441. doi:10.1108/00483480810877598

Lowden, K., Hall, S., Elliot, D., & Lewin, J. (2011). *Employers' perceptions of the employability skills of new graduates.* London: Edge Foundation.

Lunneborg, C. E. (2000). *Data analysis by resampling: Concepts and applications.* California: Duxbury Press.

MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences, 19* (4). 451-458. doi:10.1016/j.lindif.2009.03.007

Malaysian Institute of Accountants (2011). *A11 budget allocations.* Retrieved from http://www.mia.org.my/new/downloads/circularsandresources/budget/2011/A11.pdf

Malaysian Soft Skills Scale (2011). Retrieved December 26, 2011, from http://jpt.mohe.gov.my/my3s/members/index_home

Mathews, B. P., & Redman, T. (1996). Getting personal in personnel recruitment. *Employee Relations, 18*(1), 68 – 78. doi: 10.1108/01425459610110245

McCroskey, J. C. (2009). Communication apprehension: What have we learned in the last four decades. *Human Communication, 12*(2), 157 - 171.

McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology. 86*(4), 730-740. doi: 10.1037/0021-9010.86.4.730

Mertens, D. M. (2009). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative and mixed methods.* California: Sage Publication, Inc.

Miller, M. J., Woehr, D. J., & Hudspeth, N. (2001). The meaning and measurement of work ethic: construction and initial validation of a multidimensional inventory. *Journal of Vocational Behavior, 59,* 1-39. doi: 10.1006/jvbe.2001.1838

Mohamed, A. H., Shamsudin, F. M., Ismail, H. C., Teong, L. K., Saud, M. B., Zainudin, N., Mat, N. & Yaacob, N. (2004). *Pra-U STPM matrikulasi: Pengajian Perniagaan 2.* Kuala Lumpur: Pearson Longman.

MOHE (2006a). *Report: The committee to study, review and make recommendations concerning the development and direction of higher education in Malaysia: Towards excellence.* Perpustakaan Negara Malaysia Retrieved on 10 Mei 2011 from http://www.mohe.gov.my/images/Report.pdf

MOHE (2006b). *Modul pembangunan kemahiran insaniah (soft skills) untuk institusi pengajian tinggi Malaysia*: Perpustakaan Negara Malaysia.

Motor Trader (2010). *Budget 2011 allocations for training and skills programmes.* Retrieved from http://www.motortrader.com.my/JOBS/NewsHeader/Budget-2011-allocations-for-training-and-skills-pr.aspx

Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: the low-fidelity simulation. *Journal of Applied Psychology, 75*, 640-647. Retrieved on 3 April 2009 from http://www .sciencedirect.com/science/article/pii/S0021901002009846

Ngui, C. Y. K. (2005). Getting to the root of graduate unemployment. *Malaysian Business*. Retrieved on 3 April 2009 from http://findarticles .com/p/articles/mi_qn6207/is_20050916/ai_n24909278/

Nguyen, N. T., Biderman, M. D., & McDaniel, M.A. (2005). Effects of response instruction on faking a situational judgment test. *International Journal of Selection and Assessment, 13*, 250-260.

Nussbaum, B. & Johnson, J. S(2009). *Integrating conflict situations, personal characteristics, and behavior in the workplace*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University and The University of Texas.

O'Connell, M. S., Hartman, N.S., McDaniel, M.A., Grubb, W.L., III, & Lawrence, A. (2007). Incremental validity of situational judgment tests for task and contextual job performance. *International Journal of Selection and Assessment, 15*, 19-29.

Oppenheim, A. N. (2001). *Questionnaire design, interviewing and attitude measurement (New Ed.).* Great Britain: Continuum.

Peeters, H., & Lievens, F. (2005). Situational judgment tests and their predictiveness of college students' success: The influence of faking. *Educational and Psychological Measurement*, *65* (1), 70-89. doi: 10.1177/0013164404268672

Picard, O. (2007). *SJTs for GPST / GPVTS shortlisting (GP Recruitment Stage 2): Situational Judgement Tests, Professional Dilemmas*. London: ISC Medical.

Ployhart, R. E., & MacKenzie, W. I. (2010). Situational judgment tests. In S. Zedeck (Ed.), *The Handbook of Industrial and Organizational Psychology*. American Psychological Association. Washington, DC: American Psychological Association. DOI: 10.1037/12170-008

Quek, A. H. (2005). Learning for the workplace: A case study in graduate employees' generic competencies. *Journal of Workplace Learning, 17*(4), 231-242. doi 10.1108/13665620510597185.

Rasmussen, J. L. (2009). *Situational judgment test responding: best and worst or rate each response.* Unpublished master's dissertation, Texas A & M University.

Raybould, J., & Sheedy, V. (2005). Are graduates equipped with the right skills in the employability stakes? *Industrial and Commercial Training*, *37*(5), 259-263. doi 10.1108/00197850510609694.

Resampling Software (2012). Retrieved January 31, 2012, from http://www.resample.com/content/software/excel/index.shtml

Rupp, A. A. (2007). Cognitive psychometric assessment. In N. J. Salkind, (Ed.), *Encyclopedia of measurement and statistics* (vol. 1, pp. 163-164). United States of America: Sage Publications.

Salkind, N. J. (2006). *Tests & measurement for people who (think they) hate tests & measurement*. California: Sage Publications.

Saunders, V., & Zuzel, K. (2010). Evaluating employability skills: employer and student perceptions. *Bioscience Education, 15*(2), (pp. 10-25). doi: 10.3108/beej.15.2

Schafersman, S. D. (1991). An introduction to critical thinking. Retrieved on 25 January 2011 from http://smartcollegeplanning.org/wp-content/uploads/2010/03/Critical-Thinking.pdf.

Schmitt, N., & Chan, D. (2006). Situational judgment tests: Method or construct? In J. A. Weekley, & R. E. Polyhart, R. E. (Eds.), *Situational judgment tests: theory, measurement and application* (pp. 135-155). New Jersey: Lawrence Erlbaum Associates, Inc.

Singh, G. K. G., & Singh, S. K. G. (2008). Malaysia graduates' employability skills. *UNITAR E-Journal*, 4(1), 15-45.

Situational Judgment Test Marking Scheme (2012). Retrieved January 20, 2012, from http://mcqbank.co.uk/sjt-marking-scheme.

Skutch, A. F. (2004). *Moral foundations: An introduction to ethics.* Virginia: Axios Press.

Spencer, L. M., & Spencer, S. M. (1993). *Competence at work: Model for superior performance*. US: John Wiley and Son.

Urdan, T. C. (2005). *Statistics in plain English* (2nd ed.). London: Lawrence Erlbaum Associates

Wagemans, J. (2004). Cognitive psychology. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (vol. 1, pp. 351-359). Texas: Elsevier Academic Press

Weekley, J. A. & Polyhart, R. E. (2006). An introduction to situational judgment testing. In J. A. Weekley, & R. E. Polyhart (Eds.), *Situational judgment tests: theory, measurement and application* (pp. 1-12). New Jersey: Lawrence Erlbaum Associates.

Weekley, J. (2010). *Scoring situational judgment test: Does the middle matter?* Retrieved from http://www.kenexa.com/getattachment/ 7b4177fa-2842-46b2-a733-c386fb8228d2/Scoring-Situational-Judgment-Tests--Does-the-Middl.aspx.

Whetzel, D. L., McDaniel, M. A., & Nguyen, N. T. (2007). Subgroup differences in situational judgment test performance: A meta-analysis. *Human Performance*, *21*(3), 291-309.

Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review, 19*(3), 188-202. doi: 10.1016/j.hrmr.2009.03.007

Wickramasinghe, V., & Perera, L (2010). Graduates', university lecturers' and employers' perceptions towards employability skills. *Education + Training, 52*(30), 226-244. doi: 10.1108/00400911011037355

Winstead, S. A., Adams, B. L., & Sillah, M. R. (2009). Teaching the 'soft skills': A professional development curriculum to enhance the employability skills of business graduates. *American Journal of Business Education, 2*(5), 35 - 44.

Wolff, P. D. (2011). *Taking sides in a workplace conflict: A study of the influence of conflict type, hierarchy and integrity orientation on side-taking preferences.* Unpublished master's dissertation, Industrial and Organizational Psychology, University of Twente.

Workforce Skills Certification System: A Program of CASAS and LRI (2011). Retrieved December 26, 2011, from https://www.casas.org/product-overviews/assessments/workforce-skills-certification

WorkKeys: Assess Your Workforce (2012). Retrieved January 15, 2012, from http://www.act.org/workkeys/

Wye, C. K., & Lim, Y. M. (2009). Perception differential between employers and undergraduates on the importance of employability skills. *International Education Studies, 2*(1), 95 - 105.

Yuzainee, M. Y., Zaharim, A., & Omar, M. Z. (2011). *Employability skills for an entry-level engineer as seen by Malaysian employers.* Proceeding from Global Engineering Education Conference (EDUCON) 2011 (pp. 80-85). doi: 10.1109/EDUCON.2011.5773117.

Zhiwen, G., & Van der Heijden, B. I. J. M. (2008). Employability enhancement of business graduates in China: Reacting upon challenges of globalization and labour market demands. *Education+Training*, *50*(4), 289-304. doi: 10.1108/00400910810880533

Zikmund, W. G. (2003). *Business research methods* (7th ed.). Ohio: South-Western College.

BIBLIOGRAPHY

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed method approaches* (3rd ed.). Carlifornia: Sage Publications.

Hunter, D. R. (2003). Measuring general aviation pilot judgment using situational judgment technique. *International Journal of Aviation Psychology, 13*, 373-386. doi: 10.1207/S15327108IJAP1304

Lankshear, C., & Knobel, M. (2004). *A handbook for teacher research: From design to implementation*. New York: McGraw-Hill

Marshall, C., & Rossman, G. B. (2010). *Designing qualitative research* (5th ed.). California: Sage Publication.

O'Connell, M. S., Doverspike, D., Norris-Watt, C., & Hattrup, K. (2001). Predictors of organizational citizenship behavior among Mexican retail salesperson. *International Journal of Organizational Analysis, 9*, 272-280. doi: 10.1108/eb028936

Piantanida, M., & Garman, N. B. (2009). *The qualitative dissertation* (2nd ed.). California: Corwin.

Preskill, H. S., & Catsambas, T. T. (2006). *Reframing evaluation through appreciative inquiry*. Carlifornia: Sage Publication.

Sinar, E. F., Paquet, S. L., & Scott, D. R. (2002, April). *Internet versus paper selection tests: Exploring comparability issues.* Paper presented at the 17th Annual Conference of the Society for Industrial and Organizational Psychology, Toronto, Canada.

Steven, M. A., & Campion, M. J. (1999). Staffing work teams: Development and validation of a selection test for teamwork settings. *Journal of Management, 25*, 207-208. doi: 10.1177/014920639902500205

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research*. California: Sage Publication.

Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology, 49*(2), 436-458. doi:10.1037/0022-3514.49.2.436

Wagner, R. K. (1987). Tacit knowledge in everyday intelligent behavior. *Journal of Personality and Social Psychology, 52*(6), 1236-1247. doi:10.1037/0022-3514.52.6.1236

Walker, D. F. (2003). *Fundamentals of curriculum: Passion and profesionalism* (2nd ed.). New Jersey: Lawrence Erlbaum Associates.

Zhang, Y. (1999). Using the Internet for survey research: A case study. *Journal of the American Society for Information Science, 51(1),* 57–68. doi: 10.1002/(SICI)1097-4571(2000)51:1<57::AID-ASI9>3.0.CO;2-W

Appendix A

Additional Explanation of Figure 2.1 (Page 49)

The first step is to define the objectives of the instrument, clearly indicating the purpose of its development (Merten, 2009). In doing so, there are two considerations that need to be made. The first one is about the specific purpose of the proposed measurement instrument. The purpose should be consistent with the objective of the study as its implementation is to achieve the research objective. Secondly, the information to be collected on a particular attribute needs to be determined. Normally, these information can be obtained through previous researches.

Next is to identify the target respondents and the characteristic of the instrument (Dennis & Bocarnea, 2005; DeVellis, 2003; Merten, 2009). According to Merten (2009), "the relevance of the criterion of the instrument centers on the persons for whom the measurement is intended" (pp. 313). Also, the characteristics or factors related to the organization of the instrument need to be decided. Examples of the factors are the amount of time required in completing the instrument, the reading level of the respondents, the types of test items and response option, and test setting.

The third step, if available, involves revision on the existing instruments (Chisholm et al., 2000; Dennis & Bocarnea, 2005; DeVellis, 2003; Merten, 2009). Technical revisions could be made on the instrument such as the method used, the types of test item, the order of the items, and the scoring procedure.

Next is the development of the draft instrument (Chisholm et al., 2000; Dennis & Bocarnea, 2005; DeVellis, 2003; Merten, 2009). This step is the most crucial part in

the development process as this is where the realization of the instrument begins. The development process in this step can be divided into two parts that consist of the front page and the item pool. For the front page, the first thing that needs to be considered is the appropriate title for the instrument. This is because it is the first thing that will be read by respondents. The title should be motivating and consistent with the development objective. Subsequently, a short introductory paragraph needs to be included to explain the purpose of the instrument. An explanation can also be made on how and why the respondents are selected. Another important statement that should be included in this paragraph is the assurance of confidentiality. An instruction in answering the instrument and contact person should also be included.

In developing the items, if appropriate, adaptations from existing instruments with some modification can be made. However, it is also possible for the items to be developed from scratch (Dennis & Bocarnea, 2005; DeVellis, 2003; Merten, 2009). Nevertheless, there is other essential point that cannot be overlooked in developing the items i.e. the construction of the statement of the items. To put it simply, the wording should be wisely chosen to avoid any error in delivering it. Caution should be exercised on several common problems in item's wording such as complex word, double barrelled, and ambiguity. The item should be specific. However, the wording used must be simple conversational. This is to avoid confusion on the part of the respondents when answering the items (Zikmund et al., 2003).

Once the instrument has been developed, the next step is to prepare the initial draft and the implementation of pilot studies. Subsequently, the validity and the reliability of the instrument can be determined (Chisholm et al., 2000; Dennis & Bocarnea,

2005; DeVellis, 2003; Merten, 2009). Firstly, the developed instrument is recommended to be evaluated by the experts on the measured items to obtain content validity. Then, after necessary adjustment has been made based on the experts' review, the draft instrument can be tried on a small sample of targeted respondents. An extra space should be provided at the end of the instrument for the respondents to give their feedback. This is to get the general idea on the quality of the instrument. After the necessary revision has been made, the pilot study can be implemented. In running the test, the number of respondents needs to be large enough to make sure that the validity and the reliability of the instrument are able to be ascertained. If the result is not satisfactory, a correction should be done and pilot studies should be conducted again. This process will continue until the developed instrument shows the satisfactory level of reliability and validity (Dennis & Bocarnea, 2005; DeVellis, 2003; Merten, 2009).

The final step in developing an instrument is to conduct an item analysis and review the instrument (Chisholm et al., 2000; Dennis & Bocarnea, 2005; DeVellis, 2003; Merten, 2009). This step consists of an analysis using the data obtained at the previous step. Under this step, answers for each item will be reviewed in terms of their pattern to determine if the pattern suggests ambiguity or biasness. After that, if necessary, an improvement will be made. Beyond this step, the instrument is ready to be implemented for the main study.

Appendix B

List of Experts and Their Expertise

Name	Expertise	Description
Norhafezah Yusof	Communication skill	1. Senior lecturer at the Department of Communication, UUM 2. Holds a doctorate in the field of communications 3. Involved in three studies and published communication-related works 4. One of her areas of expertise is organizational communication
Nordin Kardi	Critical thinking in problem solving and professional ethics and morality	1. Contract professor at UUM 2. Holds a doctorate in adult education (Counselling & Student Development) 3. Has more than 10 research and publications related to psychology and counselling 4. Expert in counselling/student development
Norashidah Hashim	Entrepreneurial skill	1. Senior lecturer at the College of Business, UUM 2. Holds a doctorate in the field of business and entrepreneurship education 3. Involved in three studies and more than 10 publications related to entrepreneurship 4. Her area of expertise is entrepreneurship
Arsaythamby A/L Veloo	Personal quality	1. Senior lecturer at the College of Arts and Sciences, UUM 2. Holds a doctorate in the field psychometric and educational testing

Appendix C

Graduate Employability Skills Instrument

CENTER FOR TESTING, MEASUREMENT AND APPRAISAL (CeTMA)

INSTRUMEN KEBOLEHPASARAN GRADUAN

Assalamualaikum / Salam sejahtera

Kami adalah pasukan penyelidik daripada **Center for Testing, Measurement and Appraisal (CeTMA)**, Universiti Utara Malaysia, Sintok, Kedah Darul Aman. Kami sedang menjalankan penyelidikan bertajuk **Instrumen Kebolehpasaran Graduan.** Kajian ini bertujuan mengukur kemahiran kebolehpasaran graduan secara umum bagi mendapatkan gambaran awal kemahiran tersebut.

Instrumen ini mengandungi **LAPAN (8)** muka surat bercetak (Tidak termasuk muka hadapan) dan terdiri daripada **DUA (2)** bahagian iaitu:

1. Bahagian A: Ujian Penilaian Situasi
2. Bahagian B: Maklumat Responden

Segala maklumat yang diperoleh adalah **SULIT** dan akan digunakan untuk tujuan penyelidikan ini semata-mata. Kerjasama yang anda berikan sangat kami hargai. Terima kasih.

Ketua penyelidik

Prof. Madya Dr. Bidin Yatim
Pengarah
Center for Testing, Measurement, and Appraisal (CeTMA),
Universiti Utara Malaysia.
Tel (P)   :   04-9286800/01
Faks   :   04-9286802
E-mel   :   byatim@uum.edu.my
Laman web   :   http://www.cetma.uum .edu.my

BAHAGIAN A: UJIAN PENILAIAN SITUASI

Berikut merupakan **DUA BELAS (12)** soalan **SITUASI** beserta **EMPAT (4) TINDAKAN (A, B, C, dan D)**. Berdasarkan hanya kepada maklumat yang disediakan, sila isi jawapan anda di **RUANG JAWAPAN YANG DISEDIAKAN** mengikut susunan tindakan yang anda rasakan **PALING EFEKTIF** sehingga tindakan **PALING TIDAK EFEKTIF**. Tiada jawapan yang benar atau salah. Berikut adalah contoh soalan:

CONTOH SOALAN
Anda seorang pegawai yang sibuk dengan banyak urusan pentadbiran dan sepatutnya dibantu oleh rakan sepejabat yang juga seorang pegawai. Namun, kebelakangan in, beliau sering lewat ke pejabat. Apa yang perlu anda lakukan?   A. Tidak mengambil apa-apa tindakan kerana percaya bahawa lambat-laun beliau akan mendapat balasan setimpal.   B. Bekerja lebih masa untuk menyelesaikan kerja dan menganggap situasi yang dihadapi tersebut hanyalah untuk sementara.   C. Mengadu kepada ketua anda dengan harapan beliau dapat membantu menyelesaikan hal ini.   D. Berbincang dengan rakan anda untuk kenal pasti sebab kelewatannya.
BERIKUT ADALAH CONTOH JAWAPAN UNTUK SOALAN INI

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
Contoh Soalan	D	B	C	A	

Sila isi jawapan anda di **RUANG JAWAPAN YANG DISEDIAKAN** mengikut susunan yang anda rasakan **PALING EFEKTIF** sehingga **PALING TIDAK EFEKTIF**. Tiada jawapan yang benar atau salah.

1. Rakan dari Pulau Pinang melawat anda di Kuala Lumpur. Beliau menggunakan peta untuk ke tempat anda kerana tidak mahir dengan jalan di situ. Selepas kerja, anda membawa beliau bersiar-siar di bandar. Ketika berehat di sebuah restoran, seorang pelancong Inggeris mendekati beliau untuk bertanyakan jalan ke KLCC. Apakah tindakan anda?

   A. Menyerahkan kepada rakan untuk membantu pelancong tersebut.
   B. Memberi pandu arah kepada pelancong tersebut dan mengharapkan rakan untuk turut membantu menggunakan peta beliau.
   C. Anda memberi pandu arah kepada pelancong tersebut.
   D. Memberi pandu arah kepada rakan dan meminta beliau menceritakannya kepada pelancong tersebut.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
1	☐	☐	☐	☐	

2. Kumpulan anda perlu membuat laporan penilaian bagi projek yang telah dijalankan. Setiap ahli mendapat bahagian tertentu untuk disiapkan. Tiga hari sebelum pembentangan, laporan tersebut dikumpulkan. Anda dapati penulisan untuk beberapa bahagian tidak seperti yang dibincangkan. Apakah tindakan anda?

   A. Bantu mengenal pasti kesalahan dan meminta mereka tulis semula bahagian tersebut dalam masa sehari.
   B. Abaikan bahagian tersebut dan memberi tumpuan mengemas kini bahagian yang penting.
   C. Bantu ahli lain betulkan bahagian tersebut dan pada masa yang sama ajar mereka untuk menulis dengan betul.
   D. Meminta ahli lain untuk membetulkan bahagian tersebut.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
2	☐	☐	☐	☐	

Sila isi jawapan anda di **RUANG JAWAPAN YANG DISEDIAKAN** mengikut susunan yang anda rasakan **PALING EFEKTIF** sehingga **PALING TIDAK EFEKTIF**. Tiada jawapan yang benar atau salah.

1. Anda mempunyai rekod kerja yang cemerlang di syarikat A. Niat asal anda bekerja adalah untuk mengumpul modal bagi memulakan perniagaan Industri Kecil dan Sederhana (IKS). Kini, modal tersebut telah cukup. Bagaimana pun, gaji dan kedudukan dalam kerjaya anda sangat memuaskan, terjamin dan anda turut mendapat tawaran kenaikan pangkat. Apakah tindakan anda?

   A. Mengambil cuti enam bulan tanpa gaji untuk membuka dan menguruskan perniagaan. Kelak, anda akan meletak jawatan sekiranya perniagaan tersebut menguntungkan.
   B. Terus bekerja dan pada masa yang sama buka perniagaan yang tidak mengganggu tugasan anda di pejabat.
   C. Pada masa terdekat, anda akan memulakan perniagaan seperti yang dirancang dan meletak jawatan setelah perniagaan dibuka.
   D. Melupakan niat asal kerana anda telah mempunyai kerja yang dapat menjamin masa depan.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
3	☐	☐	☐	☐	

2. Anda seorang peniaga runcit yang berjaya. Namun, sejak adanya pesaing perniagaan di kawasan berdekatan, keuntungan anda kian merosot, apakah yang akan anda lakukan?

   A. Kuatkan promosi jualan kedai anda.
   B. Tawarkan hadiah dan kupon pembelian kepada pelanggan bertuah setiap minggu.
   C. Cari rakan niaga untuk meningkatkan modal bagi memajukan perniagaan.
   D. Tawarkan kepada pesaing untuk membuat usaha sama dalam perniagaan.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
4	☐	☐	☐	☐	

Sila isi jawapan anda di **RUANG JAWAPAN YANG DISEDIAKAN** mengikut susunan yang anda rasakan **PALING EFEKTIF** sehingga **PALING TIDAK EFEKTIF**. Tiada jawapan yang benar atau salah.

1. Organisasi menghantar anda ke luar Negara untuk tugasan luar selama beberapa tahun. Negara tersebut dikenal pasti mempunyai budaya dan persekitaran yang berbeza berbanding Malaysia. Apakah tindakan anda pada minggu pertama di sana?

   A. Berkenalan dengan staf yang berada di tempat kerja pada waktu kerja.
   B. Buat pemerhatian beberapa hari sebelum berkomunikasi dengan staf lain.
   C. Berinteraksi dengan staf yang berada di tempat kerja dan cuba memahami budaya kerja organisasi.
   D. Memberi tumpuan kepada kerja anda.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
5	☐	☐	☐	☐	

6. Anda ingin membeli komputer dan telah mendapatkan maklumat awal yang perlu. Setelah kenal pasti beberapa pilihan, anda ke pusat jualan komputer. Anda menjelaskan kepada penjual tentang ciri komputer yang diperlukan. Beliau mencadangkan jenama yang tiada dalam pilihan anda. Ia dikatakan sesuai dengan keperluan anda dan memiliki harga yang murah. Apakah tindakan anda?

   A. Tanya beberapa soalan kepada penjual tersebut sebelum membelinya bagi memastikan kompter berkenaan menepati ciri-ciri yang dikehendaki.
   B. Nilai semula pilihan anda dan cari maklumat tambahan berkenaan komputer tersebut serta jenama komputer yang lain sebelum membuat pembelian.
   C. Terima cadangan penjual tersebut kerana beliau lebih arif dalam hal ini.
   D. Meminta kepada penjual tersebut untuk benarkan anda cuba komputer berkenaan supaya anda akan lebih berpuas hati dengan sebarang keputusan yang dibuat.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
6	☐	☐	☐	☐	

Sila isi jawapan anda di **RUANG JAWAPAN YANG DISEDIAKAN** mengikut susunan yang anda rasakan **PALING EFEKTIF** sehingga **PALING TIDAK EFEKTIF**. Tiada jawapan yang benar atau salah.

7. Anda adalah Ketua Jawatankuasa Perasmian sebuah program universiti. Sebelum program bermula, anda dapati pengacara majlis tidak dapat hadir kerana tidak sihat. Teks majlis serta komputer riba jawatankuasa ada pada beliau. Tiada sebarang draf teks yang dapat dirujuk. Masa semakin suntuk, bagi memastikan majlis berjalan dengan lancar, apa tindakan anda?

    A. Berbekalkan atur cara program, anda menjadi pengacara majlis dan mengendalikan majlis tanpa teks.
    B. Sediakan teks ucapan ringkas dan mengambil alih tugas pengacara majlis.
    C. Mengarahkan ahli jawatankuasa yang lain untuk ambil alih tugas mengacara majlis.
    D. Mengarahkan ahli jawatankuasa yang lain untuk ambil alih tugas mengacara majlis dan anda akan menjadi pilihan terakhir untuk mengacara majlis sekiranya tiada pengganti.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
7					

8. Anda memiliki sebuah restoran. Masakan di restoran anda sentiasa mendapat pujian. Pada suatu hari, seorang pelanggan mengkritik hidangan yang disediakan tidak seperti di restoran yang lain. Cara masakan hidangan tersebut tidak betul dan rasanya tidak memuaskan. Apa reaksi anda?

    A. Terima kritikan dan akan ambil inisiatif untuk memperbaiki hidangan tersebut pada masa hadapan.
    B. Tetap dengan pendirian anda kerana setiap orang mempunyai cara masakan yang tersendiri.
    C. Terima kritikan dan akan mengkaji cadangan tersebut.
    D. Terima kritikan tersebut. Namun anda percaya setiap hidangan di restoran anda mempunyai keistimewaan yang tersendiri dan harus dikekalkan.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
8					

Sila isi jawapan anda di **RUANG JAWAPAN YANG DISEDIAKAN** mengikut susunan yang anda rasakan **PALING EFEKTIF** sehingga **PALING TIDAK EFEKTIF**. Tiada jawapan yang benar atau salah.

9. Petroleum merupakan sumber bahan api yang semakin berkurang. Andaikan Persatuan Pengguna menyarankan rakyat Malaysia supaya mengurangkan penggunaan petroleum bagi mengatasi masalah ini. Apakah pendapat anda?

A. Setuju. Penggunaan perkhidmatan awam sebagai salah satu langkah mengurangkan penggunaan petroleum lebih menjimatkan kos berbanding penggunaan kenderaan persendirian.
B. Tidak setuju. Sumber petroleum masih banyak kerana banyak lagi kawasan yang berpotensi mempunyai sumber bahan api yang belum diterokai.
C. Setuju. Penggunaan petroleum mencemarkan alam sekitar seperti isu pemanasan global.
D. Setuju. Tindakan ini dapat mengurangkan monopoli syarikat minyak dan dapat mengurangkan tekanan harga pasaran minyak.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
9	☐	☐	☐	☐	

10. Seorang staf menyertai kumpulan kerja anda. Kehadiran beliau dilihat melambatkan proses kerja kumpulan. Apakah tindakan anda?

A. Buat kerja lebih masa supaya kerja dapat disiapkan tepat pada masanya.
B. Meminta staf tersebut mempercepatkan proses kerja beliau.
C. Bantu staf tersebut dan mengingatkan beliau yang anda sentiasa sedia membantu.
D. Menyuruh staf tersebut untuk lebih luangkan masa di dalam kumpulan supaya beliau dapat membiasakan diri dengan cara kerja kumpulan dengan lebih cepat.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
10	☐	☐	☐	☐	

Sila isi jawapan anda di **RUANG JAWAPAN YANG DISEDIAKAN** mengikut susunan yang anda rasakan **PALING EFEKTIF** sehingga **PALING TIDAK EFEKTIF**. Tiada jawapan yang benar atau salah.

11. Anda ditawarkan jawatan yang lebih tinggi oleh organisasi. Jawatan itu penting dan perlu diisi segera. Tetapi anda perlu bekerja di tempat lain. Keadaan kerja untuk jawatan berkenaan menjadi persoalan kerana dalam masa dua tahun, sudah dua orang yang menyandang jawatan tersebut dan mereka berakhir dengan meletak jawatan. Apakah tindakan anda?

   A. Menolak tawaran tersebut.
   B. Menerima tawaran tersebut dan menganggapnya sebagai satu cabaran.
   C. Meminta tempoh percubaan.
   D. Menerimanya dengan redha.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
11					

12. Pensyarah anda mencadangkan supaya bilangan ahli tugasan berkumpulan ditambah dari tiga orang kepada lima orang. Apa pendapat anda berkenaan cadangan ini?

   A. Setuju, ini akan menyebabkan lebih banyak idea akan diperoleh disepanjang proses pelaksanaan tugasan.
   B. Tidak setuju, kerana ianya mungkin menyebabkan masa setiap perbincangan kumpulan bertambah.
   C. Setuju, ianya akan jimatkan masa kerana beban kerja setiap ahli akan menjadi lebih kecil.
   D. Tidak setuju, kerana kemungkinan untuk mendapat ahli yang tidak aktif lebih tinggi.

No.	TINDAKAN				Sila Kosongkan Ruang Ini
	Paling Efektif	Efektif	Tidak Efektif	Paling Tidak Efektif	
12					

BAHAGIAN B: MAKLUMAT RESPONDEN

Sila isi dan tandakan ☑ **pada tempat yang berkenaan.**

Jantina: ☐ Lelaki ☐ Perempuan

Umur: ____ Tahun

Bangsa: ☐ Melayu ☐ Cina

☐ India ☐ Lain-lain. Nyatakan: ……………

Kolej: ☐ Kolej Sastera dan Sains (CAS)

☐ Kolej Perniagaan (COB)

☐ Kolej Undang-undang, Kerajaan, dan Pengajian Antarabangsa (COLGIS)

Program pengajian: _____

Semester pengajian: ____

- SOALAN TAMAT -