

DETECTING SOCIALLY DESIRABLE RESPONSES IN PERSONALITY INVENTORY

by

PRIYALATHA GOVINDASAMY

Thesis submitted in fulfillment of the requirements
for the degree of
Master of Arts

July 2011

ACKNOWLEDGEMENT

I owe my thanks to many people who had helped in bringing my thesis to completion. Firstly, I would like to thank my supervisor, Associate Professor Dr Ong Saw Lan for her immeasurable support and guidance. Dr Ong has always facilitated me to produce work of high quality. I also would like to thank Universiti Pendidikan Sultan Idris (UPSI) for sponsoring my studies for the past two years. My sincere gratitude to Institut Kemahiran Belia Negara (IKBN), Peretak, Maktab Rendah Sains MARA (MRSM) Kuala Kubu Bharu and Sekolah Menengah Ketari, Bentong for allowing me to collect data for this study. Not to forget, thanks to my parents and family member for their continuous support and prayer for my success. Finally, I thank all individuals who has directly or indirectly helped me in getting the thesis off the ground.

TABLE OF CONTENTS

Acknowledgement	ii
Table of Contents	iii
List of Tables	vi
List of Figures	viii
<i>Abstrak</i>	ix
Abstract	xi

1. CHAPTER 1 INTRODUCTION

1.0 Introduction	1
1.1 Background of study	1
1.2 Problem statement	5
1.3 Purpose of the study	8
1.4 Research objectives	9
1.5 Research questions	10
1.6 Significance of the research	11
1.7 Limitations of the study	12
1.8 Definition of personality dimension	13

2. CHAPTER 2 LITERATURE REVIEW

2.0 Introduction	15
2.1 Personality Assessment	15
2.2 Application of Personality Assessment	17
2.3 Definition of Social Desirability	20
2.4 Social Desirability in Personality Assessment	21
2.4.1 Person as contributing factor to social desirability in personality assessment	22
2.4.2 Personality test format as contributing factor to social desirability in personality assessment	24
2.5 Methods of Reducing Socially Desirable Responding	27
2.5.1 Standard Guidelines in Minimizing Socially Desirable Responding	27
2.5.2 Forced Choice Format	29
2.5.3 Randomized Response	29
2.5.4 Equal Positive and Negative Items	30
2.6 Methods used for detecting socially desirable responding	30
2.6.1 Social Desirability Scales	31
2.6.2 Detecting social desirability from item responses	34
2.7 Theoretical Framework	40
2.7.1 Conceptual Framework	43

3. CHAPTER 3 METHODOLOGY

3.0 Introduction	44
3.1 Research Design	44
3.2 Instrument	
3.2.1 Selection of Instrument	45
3.2.2 Description of Instrument	46
3.2.3 Factorial Structure of the Instrument	49
3.4 Pilot study	55
3.5 Sampling	56
3.6 Investigation and procedure	59
3.7 Data analysis	60

4. CHAPTER 4 FINDINGS

4.0 Introduction	66
4.1 Comparing mean score of honest and socially desirable groups	
4.1.1 Comparing the mean score deviations and differences between the honest and socially desirable groups	67
4.1.2. Distribution of non-fitting response of the Rasch model	70
4.2 Identifying personality dimension that are susceptible to social desirability	
4.2.1 Factor Analysis	71
4.2.2 Test information function	76
4.2.2 (a) Test information function for the five dimensions in IPIP	77
4.3 Identifying items that are prone to socially desirable responding	80
4.3.1 Differential Item Functioning with Mantel-Haenszel Chi-square	81
4.3.2 Differential Item Functioning using Cumulative Log Odd Ratio	82
4.3.3 Differential Test Functioning	84
4.3.4 Differential Step Functioning	85

5. CHAPTER 5 CONCLUSION, DISCUSSION, IMPLICATIONS AND RECOMMENDATIONS

5.0 Introduction	89
5.1 Conclusion	89
5.1.1 Research question 1	89
5.1.2 Research question 2	90
5.1.3 Research question 3	91
5.2 Discussion	93
5.3 Implications	98
5.4 Recommendations	99

6. REFERENCES	101
----------------------	------------

7. APPENDICES

Appendix A Permission letter of IPIP: Malay version

Appendix B Permission letter: MRSM

Appendix C List of modified Malay IPIP items

Appendix D Questionnaire: Cover page of honest group

Appendix D (i) Questionnaire: Socially desirable group

Appendix EDIF output: Mantel-Haenszel approach

Appendix E (i) DIF output: Common log odd ratio

Appendix F List of item difficulties

LIST OF TABLES

Table 3.1 Classification of IPIP items accordingly to the dimension and types of items	48
Table 3.2 Total Variance Explained	50
Table 3.3 Table of Rotated Component Matrix	53
Table 3.4 Distribution of demographic factors	58
Table 3.5 Differential Step Functioning Form	65
Table 4.1 Mean and standard deviation for honest, socially desirable and combined groups	67
Table 4.2 Mean, standard deviation of five dimensions of IPIP inventory	68
Table 4.3 Score distribution for honest and socially desirable groups	69
Table 4.4 Summary of non-fitting responses according to the groups	71
Table 4.5 Total variance explained for honest and socially desirable groups	72
Table 4.6 Comparing item loading for honest and socially desirable group	74
Table 4.7 Differential Item Functioning (DIF) with Mantel-Haenszel Chi-square	81
Table 4.8 Differential Item Functioning from cumulative Log Odd Ratio	83
Table 4.9 Item difficulty parameter between honest and socially desirable groups	83

Table 4.10 Results of Differential Test Functioning (DTF) for overall and five dimensions of IPIP	84
Table 4.11 Differential Step Functioning Results for DIF items	86
Table 4.12 Summary of DIF items with DSF magnitude and interpretations	87
Table 4.13 Frequency of response categories between honest and socially groups	88

LIST OF FIGURES

Figure 2.1 Conceptual Framework	43
Figure 3.1 Scree plot to show the number of factors	51
Figure 4.2 Test information function of Openness for honest and socially desirable groups	76
Figure 4.3 Test information function of Conscientiousness for honest and socially desirable groups	77
Figure 4.4 Test information function of Extraversion for honest and socially desirable groups	77
Figure 4.5 Test information function of Agreeableness for honest and socially desirable groups	78
Figure 4.6 Test information function of Neuroticism for honest and socially desirable groups	78
Figure 4.6 Test information function of Neuroticism for honest and socially desirable groups	79

MENGESAN RESPONS IDAMAN SECARA SOSIAL DALAM INVENTORI PERSONALITI

Abstrak

Kajian ini dijalankan untuk mengesan respon idaman sosial daripada respon-respon yang berbeza. Selain itu, kajian ini juga memeriksa item dan dimensi-dimensi personaliti yang mudah terjejas kepada kebolehingan sosial. Dalam kajian eksperimental ini, 521 pelajar sekolah diuji dengan borang soal-selidik *International Personality Item Pool (IPIP)* sebanyak dua kali di bawah arahan menjawab secara jujur dan dengan kebolehingan sosial. Respon-respon dari pentadbiran pertama diklasifikasikan sebagai kumpulan jujur dan respon-respon dari pentadbiran kedua dikategorikan sebagai kumpulan kebolehingan sosial. Kepincangan skor respon dari min dan analisis keserasian digunakan untuk mengesan respon idaman sosial. Keberbezaan fungsi ujian diaplikasikan dalam mengenal pasti item-item yang cenderung untuk kebolehingan sosial. Analisa diteruskan dengan keberbezaan fungsi ujian untuk memeriksa kesan keberbezaan fungsi item pada keseluruhan ujian dan pada dimensi-dimensi personaliti yang ditaksir. Keberbezaan fungsi langkah diaplikasikan untuk mengenal pasti kategori respon yang menyumbang kepada keberbezaan fungsi item. Kajian ini juga menjalankan analisa faktor dan lengkung fungsi informasi ujian untuk memeriksa dimensi-dimensi personaliti yang rentan untuk kebolehingan sosial. Kumpulan kebolehingan sosial melaporkan taburan skor yang tinggi pada dua sisihan piawai di atas min. Bersamaan dengan itu, kumpulan kebolehingan sosial menggambarkan peratusan respon yang tidak padan dengan modal Rasch pada nilai 2.0 logits lebih tinggi

berbanding kumpulan jujur. Perbandingan struktur-struktur faktor melaporkan perbezaan di antara kumpulan jujur dan kebolehinginan sosial. Lengkok fungsi informasi ujian bagi dimensi keterbukaan menggambarkan perbezaan maklumat di antara kumpulan jujur dan kebolehinginan sosial. Enam item dikesan untuk keberbezaan fungsi item dan tiga daripada item tersebut mewakili dimensi keterbukaan. Kehadiran keberbezaan fungsi item menyebabkan perubahan dalam ujian personaliti dan dalam dimensi keterbukaan. Analisis keberbezaan fungsi langkah melaporkan kategori-kategori respon berfungsi secara berbeza mengikut item. Kajian menyimpulkan bahawa, sisihan skor yang besar dan respon yang tidak padan dalam inventori personaliti sebagai petunjuk kepada respon kebolehinginan sosial. Sementara itu, parameter kesukaran item yang rendah dan pemilihan kategori respon menggambarkan tanda-tanda kebolehinginan sosial. Lantarannya, proses penyaringan individu yang berpotensi menjawab dengan kebolehinginan sosial serta penyingkiran item-item yang cenderung ke arah kebolehinginan sosial mampu mengatasi masalah kebolehinginan sosial dalam ujian yang berkepentingan tinggi.

DETECTING SOCIALLY DESIRABLE RESPONSES IN PERSONALITY INVENTORY

Abstract

This study attempted to detect the socially desirable responses within differential responses. Besides, this study also examined items and personality dimensions that are vulnerable to social desirability. In the experimental design, a sample of 521 students was tested twice with the International Personality Item Pool (IPIP) under honest and socially desirable instructions. Responses from the first administration were classified as honest group responses and those from the second administration were grouped as socially desirable responses. The mean dispersion and fit analysis were applied in detecting socially desirable responses. Differential Item Functioning (DIF) was used in identifying items that are prone to social desirability followed by Differential Test Functioning (DTF) which examines the DIF effects to the test and to each personality dimension assessed. Differential Step Functioning (DSF) was applied to determine the contributing steps in polytomous responses to DIF. Factor analysis and item information function curves were used to examine the personality dimensions that are susceptible to social desirability. The socially desirable group reported a distinctively higher score distribution at two standard deviations above the mean. Correspondingly, the socially desirable group has a higher percentage of non-fitting responses with values more than 2.0 logits. Differences of factor structures of IPIP dimensions were found between the honest and socially desirable response groups. Test information function for the Openness dimension illustrated a differential information function between the honest and socially desirable groups. Six items were flagged for Differential Item Functioning

(DIF) and three of the items represented the Openness dimension. In addition, the presence of DIF caused changes in the test and in the Openness dimension for the socially desirable groups. The DSF analysis reported response categories function differently according to the items. The study concluded that high score and non-fitting responses in a personality inventory are indications of socially desirable responding. Meanwhile, low difficulty item parameter and endorsement of response categories are also signs of social desirability. Therefore, screening of the potential socially desirable responding individual and the elimination of items that are prone to social desirability would help to arrest the problem of social desirability in high-stake testing.

CHAPTER ONE

INTRODUCTION

1.0 Introduction

Chapter one focuses on the background of the study and problem statements which leads to the purpose of this study. Based on the purpose of the study, the research objectives and questions were stipulated. This chapter also reports the significance and limitations of this study. The chapter concludes with the definition of the five personality dimensions.

1.1 Background of the study

Testing and evaluation started as early as 2200 B.C in China in the form of civil service examinations (Aiken & Marnat, 2006). The Chinese civil service examination was used to determine whether government officials were capable in performing their tasks. This Chinese civil examination leads to the development of civil examinations in countries like Britain, France and Germany. Later in the 19th century, psychiatrists and psychologists of mental disorders developed clinical assessment techniques and tests to assess their patients (Aiken & Marnat, 2006). The great impact in the test development was when Alfered Binet constructed the first mental test called the Binet Intelligence Test. The Binet intelligence test was used to predict the scholastic achievement of an individual. The success of measuring intelligence was then applied to the Military through the Army alpha and beta tests. These military tests were constructed to measure the cognitive ability of their

recruitment candidates (Boyle, Matthews & Saklofske, 2008). Besides cognitive ability, Military test also measures the candidate's personality during the recruitment selection (Boyle et al., 2008). According to Boyle et al. (2008), the psychological testing instruments used in military recruitment gave an insight into the suitability of a recruit together with the cognitive abilities of the recruit and thus facilitated the decision making process of selecting a recruit into the Military. Among the early personality inventories were the Personal Data Sheet and Thurstone Personality Schedule.

Over the years, various personality tests were developed and among them are the Minnesota Multiphasic Personality Inventory (MMPI), California Psychological Inventory (CPI), Sixteen Personality Factor Questionnaire (16 PF) and the Revised NEO personality inventory (NEO-PI-R). Using personality to explain how human behavioural traits react within their working environment results in its importance in predicting about an individual (Detrick, Chibnall & Luebbert, 2004). It gives an idea of how individuals interact with others and their reaction towards things and ideas. The predictive nature of personality testing resulted in its wide application in the industrial, educational and in military context (Horst, 1968). Inference about an individual from the personality inventory provides extra information which would be utilized in the selection of appropriate candidates (Carrigan, 2007). Hence, the personality inventory is relied upon when screening for job applicants in employment selection (Li & Bagger, 2007).

The study by Detrick et al. (2004) proved that NEO-PI-R has predictive validity with respect to police academic performance. The predictive capacity of personality test emerges as an aid in decision making and has evolved in the Malaysian context. The significance of personality is noted as a component in employment and the educational setting. This is seen in the application of personality elements in the University Science Malaysia (USM) entrance examinations known as the Malaysian University Selection Inventory (MUnSyI) and in the entrance examinations for teacher trainees known as the Malaysian Educators Selection Inventory (MEdSI). In addition, personality elements are included in the recruitment of graduate employees into the Malaysian civil service. Even though the personality test were used under different setting, the purpose was the same; to predict an individual's attributes to aid in decision-making.

However, the use of personality tests was questioned by the Civil Rights Act of 1964 and Equal Employment Opportunity Commission (EEOC) of the United States over the validity and discriminatory issues in pre-employment (Carrigan, 2007). Doubts arose as many personal disclosures were used in the selection process and whether it contributed to any bias issues. To arrest the problem, American Psychological Association (APA), American Counselling Association (ACA), National Board of Certified Counsellor (NBCC), Standards for Educational and Psychological Testing and National Council on Measurement in Education govern and draw up the guidelines to help professionals administer the tests ethically (Drunmond & Jones, 2006). Drunmond and Jones (2006) added that the guidelines were drawn due to the impact of the personality test on society and decision makings which saw a need for a standard setting procedure in testing and assessment. Ethical measures were

considered in order to minimize any bias decision from the outcome of the personality testing.

In psychological testing, the quality of the test and ethical conduct of helping professionals were strictly monitored to reduce the biasness in decisions made. However, the test-taker's attitude and honesty are not being controlled. Studies have shown that, people like to demonstrate good impression in a psychological testing (Dunn, 2009). In cases where the test scores become a determinant for a person's future the individual has the tendency to manipulate and answer dishonestly on the test (Aiken & Marnat, 2006). Fox and Meijer (2008) also pointed out that people often respond untruthfully on personal or sensitive questions in psychological or educational assessments. The problem of untruthful responding in personality testing opens to debate over its use for the selection process (Dilchert, Ones, Viswesvaran, & Deller, 2006). Since personality items have no absolute correct answers, it makes the items easily faked (Horst, 1968). Meantime, the purpose of the test encourages a person's tendency to fake the response (Dunn, 2009). The pressure and personal need of an individual to pursue a better life would motivate the test takers to fake a test. Irrespective of the reasons, all of these would affect the validity and utility of the personality inventory (Charles, 2003).

1.2 Problem statement

Personality assessment was used to obtain data about people to reach a conclusion on clinical, legal, educational and guidance, educational and vocational selection and also for research purposes (Holt, 1971). Application of the personality instrument in the working context predicts one's personal attributes which then evaluates over the suitability of the candidate for the specific job (Fitting the bill, 2009). This is because, employee's personality traits are found to impact on their behaviour and performance at their workplace (Nek Kamal bin Yeop Yunus, 1997). The ability of the personality test to predict an applicant's capability and suitability to the workplace has resulted in taking into consideration the results of the tests for the selection procedures of employees (Kumaresan, Aizat Mohd Nasurdin & Ramayah, 2005).

The application of personality in the job screening process is found in the Public Service Department examination while recruiting a grade 41 category officer into the civil service (*Suruhanjaya Perkhidmatan Awam*, 2005). The personality of a job applicant is appraised to determine the suitability of the candidate concurrently with the position applied by the candidate into civil service (*Suruhanjaya Perkhidmatan Awam*, 2005). Meanwhile, personal attributes of students are assessed to match the corresponding choice of courses applied for in the university. The university selection tests that include personality elements are used to determine the suitability of students in their selection of course in the university (*Ujian Penetapan Khas*, 2008). University Science Malaysia (USM) conducts a test called the Malaysian University Selection Inventory (MUnSyI) to facilitate the placement of students into different courses of study in the university (*Ujian Penetapan Khas*, 2008). Besides

MUnSyI, the Malaysian Educators Selection Inventory (MEdSI) is designed specifically for filtering applicants for the teacher education training program in Malaysia (MEdSI, 2009). Applicants of teacher education programs are required to pass the MEdSI and subsequently called to attend an interview before they are chosen (Arifin Bin Hj Zainal, Asmawati Binti Desa, Hazalizah Binti Hamzah & Nachiappan, 2009). Moreover, the *Maktab Rendah Sains MARA* (MRSM) also sets *Ujian Kecenderungan Kemasukan Maktab* (UKKM) test as pre-requisites for students to be selected into MRSM (Syarat Kemasukan Ke MRSM, 2010).

The MUnSyI, MEdSI and UKKM tests are self-reported assessment instruments with the personality element included as a common domain to be assessed. In addition, a reality show called Nescafe Kisco-Start in Malaysia used the LEONARD Personality inventory to help provide additional information when selecting the 16 semi-finalist candidates (Emotional signs, 2004). This indicates the personality features can be used to predict the likelihood of a candidate and provide useful information to be used in the selection process (Fitting the bill, 2009). The predictive feature of personality assessment leads to its applications in various decision making contexts (Kline, 1976).

The findings of personality assessments in school entrance exams, university placements, job screening and competitions showed its application in high-stake testing. All these indicate the increasing reliance of personality assessments both in school in school and workplace. Therefore, similar to graduates looking for employment, students too face with high-stake tests such as UKKM, MUnSyI and MEdSI that determine their future undertakings in life. In all these tests, personality

assessment was included to gather information about them for decision making. However, the use of self-reported personality tests in competitive environments is susceptible to biased responding (Hirsh & Peterson, 2008). Moreover, when selections are meant for the best, it could place pressure on the test-taker (Nunnally, 1975). The intention to portray the best of themselves in an impressive way often leads to the manipulation of responses in a test (Lanyon & Goodstein, 1982). The motivation to distort is caused by the intended goal in a person and the implication of the implication of the test to them (Detrick & Chibnall, 2008). It is human nature for individual to portray the best impression of themselves to meet their desired goals (Iddekinge, McFarland & Raymark, 2007). Besides, an individual has the potential towards approval-seeking attitude (Leite & Beretvas, 2005). Individual's approval-seeking attitude is to meet their desired goal and to portray the required characteristics that match the purpose of the test. The approval seeking is manifested by responding to the statements in the personality tests in a socially desirable manner which becomes the concern when the results are used for making important decisions (Arthur, Woehr & Graziano, 2001).

In addition, research showed that self-reported measures often inflate an individual's score compared to rating by others (Yang, Bagby & Ryder, 2000). This is because individuals have the tendency to fake personality tests to avoid any personal disclosures (Dunn, 2009). Therefore, obtaining valid and reliable information about a person depends greatly on the cooperation from the individual (Fox & Meijer, 2008). This is because the test-taking attitude of the job applicants and students affects the actual test performance and influences the validity of selection (McCarthy & Goffin, 2003). Individuals responding in a socially desirable

manner try to present a good image rather than projecting their actual self (Horst, 1968). The socially desirable responses become a great concern to many as it reduces the validity of the personality measured (Crowne & Marlowe, 1960). This raises the question on the accuracy of prediction and becomes a threat to the validity of the personality inventory (Kroner, Mills, Yessine & Hemmati, 2004). Therefore, information obtained on individuals become unreliable and would cause a deserving individual to lose his or her opportunity in employment and education to an unqualified candidate.

1.3 Purpose of the study

Kline (1976) stated that the predictive ability of personality assessments leads to its application in various decision-making contexts. Yet, the accuracy of the personality inventory is doubtful and opened to criticism. Ellingson, Sackett and Hough (1999) mentioned that in a standard selection process, it is difficult to assess an applicant's true scores. In such circumstances, selection is done based on the observed scores which could be the true or faked responses. As stated earlier, an individual's interest in the test and his willingness determine the response patterns in the personality inventory. Therefore, it is essential to examine the score outputs as well as the scoring pattern to determine any aberrant scores. The analyses of score outputs could trace individuals who tend to respond in a socially desirable manner and determine the respondents who tend to fake or respond in a socially desirable way. Meanwhile, score output analyses would also help in extracting items that are prone to faking. The detection of deviant response patterns and items with faking possibilities could be re-evaluated prior to the selection process. Moreover, detection

of such responses would help to identify only the appropriate candidates in the selection process and reduce wrong decisions. Therefore, this study intended to examine the score dispersion of honest and socially desirable groups from the mean and person fit analysis in detecting the socially desirable responses. Differential item functioning (DIF), item difficulty parameter and differential step functioning (DSF) were used in identifying items that are prone to social desirability. Meanwhile, personality dimensions that are prone to social desirability were examined through factor analysis, test information function and differential test functioning (DTF). This would help to identify honest responses and ensure that right decisions are made about individuals to minimize errors of selecting individuals who provide socially desirable responses.

1.4 Research Objectives

In personality tests, the rightness of answers differs from one to another (Horst, 1968). This is because, individuals are different and act contradictory at times (Alex, 1975). Besides that the personality test's biased responses of individuals are caused by their intention to gain social approvals (Stocke & Hunkler, 2007). Thus, the social gaining approvals lead to difficulties in assessing an individual's personal attributes rather than thinking on how they should respond (Horst, 1968). In addition to that, personality tests with no correct answers aids in responding items in a socially desirable manner (Horst, 1968). Meantime, the transparency of the item meanings helps to clearly discriminate items which require socially desirable responses (Detrick & Chibnall, 2008). Therefore, it becomes essential to determine whether personality measures functions equivalently across different group of applicants

(Mitchelson, Wicher, LeBreton & Craig, 2009). Thus, this study was initiated to detect individuals who distort their responses in a socially desirable manner by comparing the different response patterns. This study also focused on identifying items that are prone to socially desirable responses. Finally, this study also attempted to identify dimensions in the personality inventory that are prone to socially desirable responses. Based on the above objectives, these following research questions were proposed.

1.5 Research Questions

1. To what extent do the responses of the group that distorted responses in a socially desirable manner differ from that of the honest group?
2. What are the characteristics of the items that are prone to socially desirable distortions?
3. What are the personality dimensions that are vulnerable to socially desirable responding?

1.6 Significance of the research

The application of personality measures is gaining importance in many aspects of our life (Aiken & Marnat, 2006). In tandem with this rapid growth, there are upcoming studies underlying socially desirable response effects on personality tests. Therefore, this study is an initiative to venture and explore the response patterns of individuals to detect those distorted in a socially desirable manner.

The findings from this study would help test-users to be aware of the issue of social desirability and take measures to prevent socially desirable responses. It also notifies the test-users not to rely solely on personality inventory when making inference about an individual. This is due to the vulnerability of the personality inventories.

Study also hopes to open the minds of test-takers to not accept the results of personality inventory per se as it does not measure the overall characteristics of a person. The use of statistical approaches in detecting socially desirable responses would provide evidence that even in personality testing faking can be detected. Therefore, it is hope that the detection of socially desirable responses will help to motivate individuals to provide honest responses in future. In addition, detection of socially desirable responding would help to differentiate between honest and dishonest respondents and ensure a fair chance for all in any selection process.

1.7 Limitations of the study

The accuracy of detecting socially desirable responses relies on the cooperation of the respondents. The pilot study conducted revealed that even individuals asked to respond in a socially desirable manner did not respond as instructed. This indicates that respondent participation and willingness to respond is the major limitation of this study. This is because it is difficult to monitor the respondents and direct them to respond in the way needed for the research.

Study also fails to provide an exact scenario that underlines the importance of the necessity for participants to portray their best characteristics. The instructions to respond in a socially desirable manner and scenarios presented when administering the test may not be sufficiently important to the participants. Therefore, failure in setting the appropriate environment at the experimental design would be another limitation of this study.

The choice of instrument is identified as another limitation of this study. It was difficult to obtain established instruments as they were expensive while some authors were reluctant to grant permission on the use of their inventory. Meanwhile, efforts to obtain locally developed personality instruments were also not successful. Also due to the small sample size, the data analysis using the Item Response Theory (IRT) model is limited to the 1-parameter logistic model which is run with the WINSTEP computer program that is available in the university.

1.8 Definition of Personality Dimensions

In this study, the International Personality Item Pool (IPIP) inventory was used in studying socially desirable responding. IPIP was developed based on the Big five personality model that comprised of five dimensions. The following are the five dimensions and their definitions adapted from Pervin and John (2001, p. 257).

Openness

Assesses proactive seeking and appreciation of experience for its own sake, toleration for and exploration of the unfamiliar.

Conscientiousness

Assesses the individual's degree of organization, persistence and motivation in goal-directed behaviour. Contrasts dependable, fastidious people with those who are lackadaisical and sloppy.

Extraversion

Assesses quantity and intensity of interpersonal interaction, activity level, need for stimulation and capacity for job.

Agreeableness

Assesses the quality of one's interpersonal orientation along the continuum from compassion to antagonism in thoughts, feelings and actions.

Neuroticism

Assesses the adjustments versus emotional stability. Identifies individual prone to psychological distress, unrealistic ideas, excessive cravings or urges, and maladaptive coping responses.

CHAPTER TWO

LITERATURE REVIEW

2.0 Introduction

Chapter two is a collection of critical literature reviews of social desirability and personality. This chapter begins with the definition of personality assessment and its applications in real life settings. It is followed by the definition of social desirability and its presence in personality assessments. The literature review next reviews reported methods of reducing and detecting social desirability responding. Finally, the chapter concludes with the theoretical and conceptual framework of this study.

2.1 Personality Assessment

Personality is defined as an individual's unique pattern of traits (Guilford, 1954). Alex (1975) stated that personality only refers to those traits that are assumed to be fixed in an individual. The traits convey consistent and continuing information about a person (Carver & Scheier, 2000). This is why personality is referred to as clusters of related dimensions that allow describing a person's behavior, feeling and interaction with others (Aiken & Marnat, 2006). The means of gathering and organizing information about a person with the intention of wanting to learn about a person is called personality assessment (Lanyon & Goodstein, 1982). Personality assessment comprises a series of items that describe an individual's personality (Lonnqvist, 2008). Therefore, statements in the personality assessment would be anything that can be said about a person (Horst, 1968). Meanwhile, personality

assessment presumes that there are characteristics that illustrate the differences between people and the ways to measure it precisely (Aiken, 1999).

Personality assessment takes various forms in gathering information about a person's personality. Among them are through external impressions (observation) and through self-reported measures (Carver & Scheier, 2000). Kline (1976) mentioned that personality attributes can be measured through a questionnaire, projective and objective techniques. However, most personality assessments are in the form of self-reported inventories because the items in the inventory require individuals to describe their behavior, attitude or feelings in general (Robie, Schmit, Ryan & Zickar, 2000). Due to the descriptive nature of personality assessments, it was used extensively in clinical, counseling, business, industrial, governmental, military, educational and school contexts (Aiken, 1999). The assessment and testing of personality is to explore an individual's personal attributes under various settings for the purpose of reaching a conclusion (Holt, 1971). This is because personality variables were found to contribute and correlate to an individual's life outcome events (Magnus, Viswesvaran, Deshpande & Joseph, 2006).

The first personality assessment was the Hoodsworth's Personal Data Sheet which was applied in the U.S military context (Aiken, 1999). This was then followed by assessments like Minnesota Multiphasic Personality Inventory (MMPI) applied in a clinical setting and Hogan Personality Inventory in the employment selections (Aiken, 1999). Meanwhile, the strong scientific background of the 16 Personality Factor Questionnaire (16 PF) brought its application into the industrial, organizational, counseling and clinical, research, educational and medical settings

(Cattell & Mead, 2008). On the other hand, the California Psychological Inventory (CPI) which comprises common descriptions of personality is applied in job selection procedures as the items are skill oriented and lean towards interpersonal attributes (Boer, Starkey & Hodgetts, 2008).

Despite the predictive ability of the personality inventory, it can never be adequate enough to explore an individual's personal attributes (Horst, 1968). This is because humans are complex in nature and personality inventories are merely tools in assisting the gathering of information about a person and does not measure completely a person's personal attributes in total. Incompleteness of measuring a person as a whole could also explain why individuals themselves are unaware of their attributes.

2.2 Application of Personality Assessment

Personality measures were used in personnel selections, training processes and in personal development to mould competent employees (Salgado, 2005). The personality questionnaires were used to identify individuals with traits reflecting occupational success (Kirkcaldy, 2001). Therefore, the application of personality measurement convinces that appropriate candidates were selected at the selection process (Stabile, 1997). The usefulness of personality assessment has gained its popularity and is expected to contribute to a 20% increase of online testing in future (Piotrowski & Armstrong, 2006).

Findings have reported the increasing reliance of personality testing in the uniformed forces can be seen. The necessity of appraising personality attributes was

found in the United States Army's Special Operations Target Interdiction Course (Girard & Scholtz, 2005) whereby the personal characteristics of the army officers were assessed to determine their suitability for special operations. Meanwhile, the Navy Computer Adaptive Personality Scales (NCAPS) of the United States were used in assessing individuals over their suitability to the job in the navy (Underhill & Lords, 2002). There were also studies suggesting the implementation of personality testing in the United States fire fighter's selection criteria (Younts, 2003). Besides that, the Canadian Forcers were suggesting the application of psychological screening which includes personality at the selection of Snipers (Girard & Scholtz, 2005).

Besides that, personality testing was used by NASA reviewing the suitability of an astronaut candidate and adding them in their selection of astronauts (Musson, Sandal & Helmreich, 2004). In competitive environments, more information about a person is required in making decisions (Oswald, Schmit, Kim, Ramsay & Gillespie, 2004). Therefore, non-cognitive elements such as personality are used to predict individual's suitability during the selection of students into the higher learning (Oswald et al., 2004).

Psychological measures are also used in diagnosing the problems of an individual in schools and institutions and to apply interventions for problematic students (Wrobel & Lachar, 1998). Similarly, the Royal Australian Air Force (RAAF) uses personality measures in diagnosing the student's problems and assisting them in overcoming their difficulties during their training (Elliot, 1997).

In Malaysia, the Public Services Department (PSD) is the largest employer of the nation. PSD conducts general assessments in the process of recruiting a graduate candidate for Grade 41 in the civil service (*Suruhanjaya Perkhidmatan Awam*, 2005). This assessment consists of two sections where section I comprises items that embrace general knowledge of Malaysia and its surroundings, problem solving ability, writing and comprehension skills in *Bahasa Malaysia* and the English language, Section II assesses the non-cognitive aspects, that is on personality measurement. Personality is appraised to determine the suitability of a candidate for the job that the candidate is applying. This highlights the importance of personal attributes in recruitment in addition to cognitive abilities. Even after the recruitment, the Public Services Department still enhances an employee's personality in order to ensure a success-oriented culture in the civil service (*Surat Pekeliling Perkhidmatan*, 2009).

Personality assessments are also being used to determine the suitability of students as regards to their choice of course when applying for a place in the university (*Ujian Penetapan Khas*, 2008). University Science Malaysia (USM) administered the Malaysian University Selection Inventory (MUnSyI) to facilitate the placement of students into the appropriate courses of study (*Ujian Penetapan Khas*, 2008). MUnSyI encompasses five distinctive domains including career interest, personality, integrity, emotional intelligence and patriotism. The MUnSyI self-reported measure outcomes are used to place students into different courses in the university.

Besides the MUnSyI test, the Malaysian Educators Selection Inventory (MEdSI) is another test administered to candidates applying for teacher degree programs in Malaysia (MEdSI, 2009). The self-reported MEdSI measures four domains which are career interest, personality, integrity value and emotional intelligence. MEdSI is very specific in nature because it is designed particularly for teacher candidates. Therefore, the four domains would assess candidates over the qualities required in the teaching profession. Nonetheless, candidates of MEdSI are required to pass the MEdSI test in order to proceed to the interview session before they are recruited into the teaching program (Arifin Bin HjZainal et al., 2009).

2.3 Definition of Social Desirability

Social desirability is termed as the tendency to respond to self-reported items in a way that presents the individual in a good light rather than to respond in an accurate and truthful manner (Holtgraves, 2004). This has led to a systematic responding to items on some basis other than the specific item content (Robinson, Shaver & Wrightsman, 1991). Socially desirable responding was described as the way individuals project themselves in socially desirable image (Cervellione, Lee & Bonanno, 2009). Richman, Kiesler, Weisband and Drasgow (1999) defined social desirability as respondent's propensity in manipulating responses in a socially desirable manner under different test conditions, mode and administration. Therefore, social desirability can be concluded as the tendency for individuals to distort their responses according to the test conditions with the intention to present themselves in socially desirable and acceptable images.

Social desirability is associated to impression management. The term, impression management is defined as consciously and intentionally distorting responses towards positively desirable ways (Charles, 2003). Impression management was also known as motivational distortion which takes the form of faking good that refers to socially desirable responding (Hakstian & Ng, 2005). Meanwhile, impression management was also described as one's purposeful response to create the most positive social image (Robinson et al., 1991). Stocke and Hunkler (2007) defined impression management as respondents giving biased answers with the intention of obtaining social approval from others. In addition, impression management was classified as a voluntary projection of individuals on how others want to perceive them in a positive way (Salgado, 2005). Impression management's underlying concepts of intentional distortion of responses to portray socially desirable image for other's acceptance does show similarity between the social desirability and impression management terms. Although the terms differ, the meaning of social desirability and impression management are the same.

2.4 Social Desirability in Personality Assessment

The literature review found that the opportunity to fake, personal characteristics of individuals and situational factors contribute to faking (Magnus et al., 2006). Research on person and personality test format that contribute to socially desirable responding are discussed in detail under each following subtopics.

2.4.1 Person as a contributing factor to social desirability in personality assessment.

Morgeson, Campion, Diboye, Holleabach and Schmitt (2007) in their study reported faking as a result of person and situation interaction. This means the degree of faking differs between individuals and the situational demands, whereby the situations of the job applicants provoke the applicant's motivation to distort their responses. Indeed, job applicants were found to be more motivated to fake their responses compared to the incumbents of the job (Day, 2008). In a study by Harvey, Wilson and Hansen (2005), instructions to fake the personality test to the Troopers officers found score elevations compared to their responses in a honest condition. In addition, when test users are viewed in applicant condition, they tend to elevate their scores significantly compared to the honest condition (Harvey et al., 2005).

According to Hakstian and Ng (2005), job applicants distort the responses to increase the likelihood of getting the job they have applied for and the distortion occurring in employment is referred to as employment related motivational distortions. This is because respondents manipulate their answers to gain social approval from others (Stocke & Hunkler, 2007). However, socially desirable responding is linked to the test taking situations (Kroner et al., 2004). In this instance the test administration and implications influence the occurrence of socially desirable responding. Studies have found the discrepancy between the observed and true nature of a respondent in a motivating setting as intentional distorting or as faking (Dilchert et al., 2006).

Meanwhile in Edith, Stoop and Meijer's (2000) study, it was mentioned that a person's ability can be invalidated due to the familiarity of the questions as well as

guesses of answers and the pre-knowledge of some items. Moreover, the familiarity of items could easily determine the social values placed on the scale item which leads to the likelihood of socially desirable responding (Dilchert et al., 2006). Therefore, the variety of personality testing found online and in magazines exposes the test-taker over the likelihood of the items in personality testing. Moreover, free and accessible personality inventories attract more individuals to test and infer their characteristics. The great exposure on personality inventories lets the test-takers map out items that would respond to a particular construct in a personality testing.

In the study by Magnus et al. (2006), relationship was examined between social desirability with emotional intelligence, over claiming and self-esteem variables. Their study found that lower self-esteem individuals were more likely to be engaged with socially desirable responding. While, individuals with greater emotional intelligence also had a tendency to socially desirable responding. It is believed that lower self-esteem individuals try to respond in a socially desirable manner with the intention of gaining social approval from others. However, emotionally intelligent persons are well adjusted individuals and their socially desirable responding could be their actual self-description.

The study by Iddekinge et al. (2007) mentioned that impression management was found in the form of verbal statements, non-verbal behaviors and modification of appearance during an interview session. Explanations of past experience and competence in Behavioral Descriptive Interview (BDI) approach in interviews are more likely to use self-promoting impression management (IM). Even though the structure of the interview and the interviewer hints of the interviewee in portraying

the best of interviewees, the degree of impression management is moderated by the situational factors.

Meantime, Robie, Tuzinski and Bly's (2006) survey on assessor's beliefs and practices in recruitment related to faking revealed that candidates desire to present themselves in the best light as the reason behind their faking. In their study, they reported that the recruitment assessor believed faking is a threat to the validity of a personality inventory yet agreed that it does not affect the overall assessment of an individual. This is due to the test content and the outcomes of the test that influence the reactions to the personality tests. As a conclusion, the test implications and consequences motivate an individual to distort the response in a socially desirable manner with the help of the past experiences of the individual.

2.4.2 Personality test format as contributing factor to social desirability in personality assessment.

Measures of non-cognitive variables are found to be vulnerable to response biases (Rennie, 1982). This is because research has demonstrated evidence of faking of non-cognitive self-reported measures (Alliger & Dwight, 2000). According to Kroner et al. (2004), self-reported inventories are vulnerable to alternative interpretations which then become a threat to the validity of the inventories. A study by Yang et al. (2000) found that a self-reported inventory has higher mean differences compared with others rating on individuals using the same inventory. It is supported that the mean for responses changes according to the faking instructions given to the respondents (Viswesvaran & Ones, 1999). Viswesvaran and Ones (1999) in their study also reported that respondents are able to elevate their responses to half

a standard deviation when asked to fake. This is because individuals have the tendency to endorse items on interest highly to match the purpose of the test.

The study by Yang et al. (2000) found that NEO PI inventory was susceptible to response bias. The response bias in NEO PI was examined and found that those respondents who were asked to fake good scored lower on Neurotic and higher on Agreeableness and Conscientiousness dimensions. Inversely, the fake bad respondents scored low in Agreeableness and Conscientiousness and higher in Neurotic dimension. This finding shows the ability of the respondents to identify the items representing different dimensions in a personality inventory. It could be due to the familiarity of extraversion and neuroticism dimension in personality which makes the items representing the dimension recognizable and therefore easily manipulated according to the situational demands (Kirkcaldy, 2001).

Meantime, Steffan, Kroner and Morgan (2007) in their clinical setting found that prior knowledge of a symptom helps individuals in dissimulating malingering effects. Prior knowledge of the symptom would help to determine the desirable and undesirable symptoms of a specific medical condition. Advancement over the knowledge then would guide them to respond to items assuming to be the best answers which would eliminate them from any medical conditions. Besides that, the transparent meaning of items clearly indicates the most positive answers to the candidate which then enables the candidate to fake the answer easily (Morgeson et al., 2007). Holtgraves (2004) mentioned that the items ambiguity also contributes to the problem of socially desirable responses. It was stated that items with high ambiguity makes the test less stable under repeated measures (Ferrando, Lorenzo &