Personality Assessment and Ethnicity:

A New Zealand Study

A thesis presented in partial fulfilment of
the requirements for the degree of
Master of Arts in Psychology at Massey University, Palmerston North
New Zealand.

By Simon Cox
Department of Psychology
Massey University

2008

Abstract

Ethnic status is one of the most protected demographic groups in terms of test bias and discrimination in personnel selection, as such bias breaches many laws, morals, and ethical procedures. To date there has only been two published studies that have used New Zealand relevant ethnic groups when analysing whether personality measures used in pre-employment settings exhibit different mean scores. The present thesis performed a systematic evaluation on the impact of ethnicity on personality traits. The study examined the impact of ethnicity on NZ Army Officer applicant personality assessment scores measured by the EPQ-R and GPP-I. Four ethnic groups (NZ Europeans, Mâori, Pasifika, and Asian) were analysed for mean trait score differences among ethnic groups, the stability of these differences across different personality inventories and models, the variance of personality traits, the impact of ethnicity on age and gender relationships with personality traits, and how these differences were related to employment selection outcomes.

The analyses revealed that ethnicity did have some impact on mean personality assessment scores used in the study. However, these were mostly small differences among ethnic groups. In addition, most of these differences found on the EPQ-R and the GPP-I were not consistent across inventories and models. There were no significant variance differences found on personality traits among ethnic groups. While initial analysis suggested that ethnicity did not influence the relationship between age and gender on personality assessment, further examination suggested that the relationship between gender and personality was impacted on by ethnicity. There were reassuring results found for New Zealand psychologists and HR specialists, as only two of the twelve traits analysed showed moderate differences on traits that were related to selection outcomes. However, for the NZ Army OSB selection process the findings in the present study indicate that the Lie scale on the EPQ-R and the Vigor trait on the GPP-I may need to be interpreted with caution for Asian and Pasifika groups. These results are discussed in terms of implications for personality theory, measurement, and the direction of future research.

Acknowledgements

To all those persons who, in special, and often quite different ways, have assisted me in this research, I extend my grateful thanks. In particular I am especially grateful to my supervisor, Dr Gus Haberman, who has offered continued support, inspiration and encouragement throughout this process. I would also like to thank the New Zealand Army Psychology group, particularly Major Helen Horn and Major Stephen Kearney, for allowing me access to New Zealand Army data and setting aside their time to help with research agendas. Finally, I would like to express my tremendous gratitude to both my mother and father and family as a whole who have provided me with great support, guidance, wisdom and love during the duration of this thesis.

Table of Contents

List of Tables

List of Figures

1. Research Overview

The use of psychometric tests is ubiquitous in New Zealand personnel selection procedures today. This means they are used for a purpose that can affect individual's working lives dramatically, for good or for bad. For example, psychometric tests are now used in employment settings ranging from selection decisions, training, and health and safety issues. Because of the powerful presence psychometric testing has in employment settings, it is important and necessary for researchers to explore and examine potential issues involved in psychological testing. This requires full and fair appraisements on the advantages and drawbacks of various forms of psychological tests (Salgado & Fruyt, 2005).

One of the major forms of psychometric assessment is personality assessment. In recent years, personality assessment has increasingly been used by New Zealand organizations and consulting firms (Taylor, Keelty & Mcdonnell, 2002). Personality assessments are used so readily in New Zealand that, in a survey of personnel selection methods across 18 countries, Ryan, Mcfarland, Baron, and Page (1999) identified New Zealand organizations as being among the highest users of personality assessments.

Since formal personality assessments were first used for selection purposes in 1917, for the screening of World War I draftees, scepticism surrounding the validity and utility of this form of assessment has existed (Cervone & Pervin, 2008). However, beginning in the 1990s support for personality tests, in personnel selection, increased with renewed optimism and vigour, as meta-analytic research demonstrated the association between personality constructs and job performance (Barrick & Mount, 1991; Hogan & Holland, 2003; Judge, Bono, Erez, Locke, & Thoresen, 2005; Judge & Ilies, 2002; Tett, Jackson, & Rothstein, 1991). Since the development of this research suggesting that personality measures are valid predictors of human behaviour the field of personality psychology has continued to aggrandize, as research is now conducted in a variety of countries and cultural settings. Over the last two decades research on personality has begun to examine more complex forms of validity with some research now beginning to examine the controversy involving whether

personality measures, derived from the English language and Western culture, can be interpreted or modelled in the same way across cultures.

Recent research in this area suggests that the personality trait structure of the FFM (Paunonen & Ashton, 1998) and Hans Eysenck's Three Factor Model (e.g. Barrett & Eysenck, 1984) are universal across cultures. These findings that personality structures can be replicated across diverse ethnic groups and cultures emphasises the robust and universal nature of personality traits. However, cross-cultural validity of models of personality does not necessarily mean that trait behaviours are equivalent across cultures (Eysenck, Barrett, & Barnes, 1992).

The comparison of mean personality trait scores across ethnic groups, despite being intrinsically interesting, may provide awareness into more complex forms of test validity as well as increase understanding on the potential adverse impact of personality tests in personnel selection settings. The importance of research on this more complex form of validity testing is threefold. Firstly, ethical guidelines for psychologists suggests that accurate assessment of ethnicity/cultural differences in personality are needed to help psychologists prevent or correct "practices that are unjustly discriminatory" as required by their ethical principles (section 1.2.1, New Zealand Psychological Society, 2002).

Secondly, present opinion states that for a test to be fair when used for selection purposes, it must conform with equal opportunity laws and codes of conduct, which prohibits discrimination in employment on the grounds of race, colour, religion, national origin, age or gender (Employment Relations Act, 2000). Thus, with general practice for employee selection having practitioners matching personality test scores (i.e. the personality profile) with job characteristics derived from a job description (Salgado & Fruyt, 2005), different personality scores for different ethnic groups may lead to adverse impact for minority groups. Finally, if different personality assessments have the potential to result in fewer applicants from minority ethnic groups being hired by an organization, there may be a negative impact on organization performance using personality assessment for hiring (e.g. De Dreu & West, 2001).

Like most areas of psychology the research on personality differences among ethnic groups has predominantly been conducted overseas, with most of this research being conducted in the United States and more recently in Britain, East Asia, and South Africa. The result of this has meant that there is almost no research focusing on ethnic groups that are specific and relevant to New Zealand. Further, in a review of the available literature on this topic there were only two published articles found that have examined New Zealand groups (Guenole & Chernyshenko, 2005; Packman, Brown, Englert, Sisarich, & Bauer, 2005).

The two studies conducted in New Zealand are a significant first step for research on the relationship between ethnic groups and personality assessment. However, the two studies do have significant shortcomings that require further research to be conducted on this issue. Firstly, because the nature of Guenole and Chernyshenko (2005) was not to explore ethnic mean trait differences but rather the suitability of the FFM model for selection in New Zealand, the design of the study had a significant methodological limitation in investigating ethnic differences. The study suffered from range restriction with the sample group only consisting of employees in various occupations. Secondly, neither study analysed the stability of the differences between ethnic groups on personality assessment scores, as only one set of data was obtained from the personality inventories used in the respective studies. This is an important shortcoming as it is well understood that psychological constructs such as intelligence, anxiety, and personality traits such as Extraversion have manifold and arguably value implications that may or may not be sustainable across associated measures. Thirdly, in relation to the second limitation, the two previous studies are based solely on personality inventories empirically related to the FFM (i.e. FFM), thus, making it difficult to generalise to alternative personality assessments that are based on alternative personality constructs (see section 2.1.2 page19).

In addition to these shortcomings, previous New Zealand studies on ethnic differences in personality assessment have only focused on mean differences. However, there are at least two other ways that the ethnic/cultural variable can be associated with personality; (i) ethnicity could influence the variance of one or more personality traits, (ii) ethnicity might moderate the relation between a pair of measures.

1.2 The Current Study

The purpose of the proposed research is to collect further data on the relationship between ethnicity and personality, and gain more practical information regarding the implications of this relationship to personnel selection strategies. There are six major aims in the current study.

To:

(i) Analyse New Zealand data to find potential relationships between personality and ethnicity. This includes comparing ethnic groups along relevant personality variables (e.g. by means of two-tailed *t* tests utilising the means as well as effect size differences and standard deviations of ethnicity defined subsets among the participants), as well as an analysis of the impact of ethnic status on gender - and age - relevant patterns in personality assessment.

(ii) Provide further insight as to whether statistical comparisons between ethnic groups reflect "real" similarities and differences.

(iii) Provide data relevant to New Zealand ethnic groups for other models besides the FFM that are currently used in personality research and personnel selection.

(iv) Check whether any relationships between personality and ethnicity can be shown to be pervasive, irrespective of choice of a systemic model.

(v) Examine whether differences observed on personality assessment scores significantly impacts on employee selection outcomes.

(vi) Provide further data on the relationship between personality and ethnicity for New Zealand sample groups.

In the absence of a coherent body of literature on this topic and comprehensive meta-analytic studies it would be perilous to make prediction about potential findings, as there is no protection from the potential relations between personality variables and ethnicity. Therefore, the current study provides an exploratory approach.

1.3 Overview

The following chapters review the literature regarding the relationship between ethnicity and personality. To allow the reader full insight into the issues behind personality and ethnicity, chapter 2 discusses personality theory examining how personality assessment defines personality and reviews validity and utility issues that have allowed for the exploration of ethnicity in relation to personality. Chapter 3 then reviews research examining the relationship between ethnicity and personality and also outlines the importance and necessity of such research. Chapter 4 then outlines the design and aims of the current research.

2. Personality Assessment and its Utility in Personnel Selection

2.1 Personality Assessment: An Overview

Personality psychology is concerned with the nature of human nature – in selection settings it is a view of human performance from a broad contour. In other words, personality is concerned with systematically investigating individual differences and the organized functioning of the person as a whole (Pervin, 2003). The objective assessment of personality used to predict an individual's future behaviour has frequentely been criticised with many scientific findings questioning the validity and utility of personality assessment (e.g. Carlson, 1971; Fisk, 1974; Mischel, 1968, 1973). This has historically led to the perception that personality measures are easily faked, lack validity, and are therefore unsuited for preemployment screening purposes. However, contemporary research indicates that such perceptions have a small influence on the predictive power of personality assessment for employment settings, thus, allowing the exploration of more technical forms of validity e.g. ethnic, gender, and age influences on the psychometric properties of assessment. The present chapter briefly reviews the validity and utility issues that personality assessment has historically faced and then discusses two of the most proponent models of personality which have assisted with the establishment of predictive power for personality assessment. These models are the Five Factor Model (FFM) and Eysenck's Three Factor Model. But first a brief discussion on the theory behind objective personality assessment.

2.1.1 Theoretical Bases of Personality Assessment

"Why is it that while all Greece lies under the same sky and all the Greeks are educated alike, it has befallen us to have characters variously constituted?" (cited in Eysenck, 1982, p. 1). Theophrastus a Greek philosopher raised this question about human behaviour two thousand years ago showing us that the curiosity behind

personality is not a new phenomenon. Still, personality research has struggled with its progress in answering this question, as the complex and diverse nature of personality has led to vast disparities of opinion on what is meant by the term personality, and consequently this has also lead to disparities in methods used to develop understanding of human behaviour.

Historically there have been many different approaches to the study of personality. The earlier years of scientific research on personality saw the rise of the psychodynamic approach. This approach suggested that the goal of human behaviour is to reduce tension or release energy that is a result of energy flow in the human brain. The theory came to great prominence during the first half of the twentieth century, as personality measures based on psychodynamic theory (projective tests) were administered to literally millions of individuals throughout the world (Lilienfield, Wood, & Garb, 2000).

During the prominence of psychodynamic approach there were key developments that took place in psychometric theory that sanctioned personality theorists to administer more systematic data collection and perform more technical statistical data analysis (i.e. correlation and factor analysis). With these developments in measurement the field of personality began to take a more rational and objective approach to the theory of personality. This required the measurement of personality to be subjected to questions regarding their reliability and validity and eventually led to projective tests used in psychodynamic theory being subjected these measurement standards. These new standards led to many concluding that the projective tests used for measurement in psychodynamic theory commonly did not work (Dawes, 1994; Lilienfield et al. 2000; Rorer, 1990). This finding further led to a large reduction in the administration of projective tests, and currently their use is opposed by many psychological associations throughout the world (e.g. American Psychological Association).

During the decline of the psychodynamic approach to personality research there have been many other personality theories that attempted to provide explanation for human behaviour e.g. Behaviourist theories, Cognitive theories, Humanistic theories, Biopsychological theories. One of the approaches that is most commonly used for personality assessment in contemporary psychology is the trait approach.

Trait Theory – The Concept of a Trait

The trait approach to personality theory views personality as consisting of a variety of dimensions that individuals can differentiate on to varying degrees (Hough & Furnham, 2003). Trait psychologists view traits to be the major unit of personality, as many feel that it has helped provide stability and consistency to the study of personality (e.g. Buss, 1989; Brody, 1988; Goldberg, 1993).

Trait theory involves an objective and rational approach to the study of personality. Although the starting point of trait theory can be traced to the beginning of personality theory, the study of personality traits using scientific methods only began within the past century (Cervone & Pervin, 2008). One of the major reasons for this late development was the requirement for technological developments that allowed personality theorists to systematically collect data, use technical statistical data analysis, and the development of testable theories (Hogan, Johnson, & Briggs, 1997). Of key importance to these developments was the development of correlation and factor analysis (Gorsuch, 1983) and later Thurstone's (1947) multiple factor analysis. After these developments personality research saw an influx of influential studies using the trait approach (e.g. Allport, 1937; Galton, 1884; Spearman, 1904; Webb, 1915).

Perhaps due to the relatively infant stages of the scientific study of personality traits the field of trait theory is very different from other areas of personality. Unlike other areas of personality, such as Psychodynamic and Phenomenological traditions, there has been no one influential figure in trait theories of personality (Cervone & Pervin, 2008). This has led to the development of a variety of contributions that are distinct and diverse (e.g. The Five Factor Model, Cattell's 16 Personality Factors, Esyenck's Three Factor Model). With the disparity of conceptions surrounding trait theory some may come to the conclusion that an attempt to define, and therefore, research the area comprehensively is beyond any one study. However, although a universally agreed definition is perhaps beyond ones grasp a relatively broad definition that refers to the shared assumptions and themes is possible.

With all the disparity and diversity amongst trait psychologists there are some general assumptions and themes that they generally agree upon. Firstly, it is commonly agreed that personality traits refer to distinct and consistent patterns in the way individuals behave, feel, and think. For example, if one describes an individual as being aggressive, one would expect them to portray this aggressive trait across a variety of situations. Secondly, related to the first assumption, it is generally agreed that people carry broad dispositions to respond in diacritic ways. Thirdly, there is direct concurrency between a person performing a trait related action and his or her possession of the corresponding trait. Finally, human behaviour and personality can be organized into a hierarchical structure (e.g. The Five Factor Model, Eysenck's Three Factor model).

With the knowledge of these shared assumptions we can now look at a definition. In a review across the literature it appears that the best definition that incorporates these assumptions is one produced by Allport who is considered to be one of the most influential figures in personality trait research, as his work has been the main impetuse for a substantial amount of trait research in the 20th and 21st century (Cervone & Pervin, 2008). Allport defines personality traits as "generalized and personalized determining tendencies – consistent and stable modes of an individual's adjustment to his environment" (Allport & Odbert, 1936, p. 26). Allport's definition stresses two connotations regarding traits; (i) traits describe a regularity in human behaviour, and (ii) traits describe differences in human behaviour.

Contemporary research advocates that there are two forms of traditionally distinguishable traits (Mackinnon, 1951); those traits that can be directly observed by others (behavioural traits) and those traits that must be inferred (emotional and cognitive traits). Meehl (1956) described these traits as *phenotype* and *genotype* respectively. There are two common assumptions that are associated with the two forms of traits. Firstly, outer traits are assumed to be descriptions of behaviour that need to be explained, while inner traits are assumed to involve the causes and reasoning that explain the outer traits. Secondly, observers are assumed to have privileged access to outer traits while actors are assumed to have privileged access to inner traits. The second assumption is relevant to the validity of observer ratings versus self rating personality questionnaires.

Both forms of personality traits are often viewed to be present along a continuum. Allport (1961) suggests that common personality traits conform to the normal bell-shaped curve in which some people display extremely high and low scores on a measured trait but most people fall near the middle of the continuum. Research generally suggests that traits configure themselves to this conceptual pattern when measured in large heterogenous samples (Costa & McCrae, 1988). Many trait theorists point to the nature of traits for following a normal distribution as the prime difference between using traits as opposed to *types* for the unit of analysis in personality. However, this is perhaps one of the more negligent areas of trait research.

The type concept has often been criticised for postulating categorical distinct groups of people. Yet although this was the view of early type advocates such as Immanuel Kant, this is not the position that has been preserved. This is best demonstrated by remarks given by the most influential theorist in contemporary type theory Carl Jung "there can never be a pure type in the sense that the one mechanism is completely dominant to the exclusion of the other" (Jung, 1921, p. 21). Rather than discontinuity being the major difference between the type and trait approaches, some suggest it is rather the level of description (Eysenck & Eysenck, 1985). This view suggests that a type can be thought of as a correlated group of traits, just as a trait is thought of as a group of correlated behavioural acts or action tendencies. According to this view, then, the difference between traits and types is not the distribution of these units of analysis, but rather the level of analysis.

Because of it's objective nature, trait theory is largely dominated by measurement. Moreover, trait theoreticians consider that the ability to measure psychological traits, reliably and validly, is the critical step in building the science of personality (Cervone & Pervin, 2008). Perhaps the popularity of the trait approach to personality research is best illustrated by the vast amount of objective personality assessments in various forms that have manifested in the realm of personality research. In contemporary psychology there are literally hundreds perhaps thousands of personality assessments designed to measure individual differences for a variety of purposes (e.g. Clinical practice, Personnel Selection). Thus, it is perhaps of no surprise that much of the controversy and scepticism surrounding personality assessments has been largely

based on reliability, validity, and utility issues. The next section of this chapter will briefly review the history of personality assessment with a major focus on these issues.

2.1.2 Personality Assessment in Psychometrics

Historical Overview

The development of self-report objective personality measures were first attempted at the turn of the twentieth century. The evolution of this form of testing was quickly induced by the intense need to mobilize large military forces in WWI (Evers, Anderson, & Voskuiijl, 2005). After WWI the development of personality measures continued to flourish, as research led to the development of multi-trait tests (e.g. Berneuter, 1931, Personality Inventory) that split the whole person into decontextualized dispositional constructs (Hogan, Johnson, & Briggs, 1997).

The development of a vast array of personality measures led to large amounts of criticism and scepticism surrounding both issues of measurement and theory. This early scepticism was tempered by a degree of optimism that great knowledge was just around the corner (e.g. Bronfenbrenner, 1953; McClelland, 1951; MacKinnon, 1951; Nuttin, 1955; Sears, 1950). Nevertheless, signs of discontent continued to rise throughout the 1950s, as frustration continued to culminate over contradictory empirical results, and an overall lack of coherence. By the late 1960s, personality measures and the field of personality psychology was beginning to be thought of as a science that yielded results that were inconsistent and pointless (e.g. Adelson, 1969; Sechrest, 1976). Moreover, in a comprehensive review of the research literature of that time Guion and Gottier (1965) concluded that "…it is difficult to advocate, with a clear conscience, the use of personality measures in most situations as a basis for making employment decisions' (p.166).

The scepticism over the use of objective personality assessments was particularly high during the 1970s, as scientific publications casting doubt on the theoretical assumptions behind identifying types of individuals continued to accumulate. There

were a number of influential critiques of personality assessment that contributed to this doubt (e.g. Carlson, 1971; Fisk, 1974; Mischel, 1968, 1973). Carlson and Fisk both heavily criticised the field of personality psychology on how knowledge was being accumulated. For example, Carlson advocated that most research in personality psychology was measuring only one dimension of a person, thus, not fulfilling the approach advised by Allport and Murray concerning personality's commitment to the investigation of the whole person.

While there were many scholars who contributed to this debate, the empirically based critic, Mischel (1968, 1973) was perhaps one that held the most influence. In *Personality and Assessment* (1968), Mischel claimed that personality dispositions, often measured via paper-and-pencil tests and questionnaires, accounted for very little of the variance in human behaviour and that the usefulness of such measures up to that point had been seriously overestimated. Mischel argued that personality or personality traits were only relevant for the prediction of behaviour when the trait or dimension of personality is relevant, salient, or important for the person in a situation. Thus, he suggested that there was very little cross-situational generality in behaviour. Instead human action was better interpreted by situational specific factors. Mischel (1968) provided evidence to back up his claims with findings showing that personality variables had low cross-situational or temporal consistency (less than .30), and did not correlate well with behavioural outcomes.

The critiques of Carlson, Fiske, and Mischel were met with a barrage of counter criticism (Alker, 1972; Bowers, 1973; Epstien, 1979; Snyder, 1983; West, 1983). This led to psychology journals filled with lengthy debate over the relevance of Carlson, Fiske, and Mischel's claims. This debate prompted an effort to improve the predictive power of personality measures, with certain personality psychologists championing (a) moderator variables and (b) aggregation as the reason for inconsistent and low cross-situational or temporal consistencies. For example, with respect to the first, Bem and Allen (1974) provided research suggesting that predictions of behaviour can be enhanced when personality measures are used together with assessments to the extent certain traits are likely to be relevant, salient, or important for the person (e.g. Job Analysis in a selection setting). In reference to the latter an accumulation of research has been developed suggesting that when functionally similar acts are aggregated over

time and across situations, reliability is enhanced and personality measures are able to obtain a better prediction of behaviour (Epstein, 1979, 1984; Snyder, 1983).

The empirical and theoretical debate over the accuracy of personality measures in predicting behaviour continued into the 1980s. However, by the late 1980s, after technological developments allowing more technical statistical analysis (Goodstein & Lanyon, 1999), revisions on the structure of personality (e.g. Five Factor Model, The Three Factor Model), more contemporary and better validated assessments, and the publication of personality inventories deliberately targeted to occupational settings, researchers and HR specialists became less sceptical about using personality data for predictors of job performance.

Following the gain in confidence in personality measures in the late 1980s, the 1990s saw renewed optimism and vigour for personality measures (Hogan & Jones, 1985; Maddi, 1984; West, 1983). The confidence in personality measures to predict human behaviour became so great during the 90s they became common measures in personnel selection (McAdams, 1997). The reason behind this renewed confidence was due to a wave of research that began to demonstrate comprehensively an association between personality constructs and job performance. There were two meta-analytic studies on personality and job performance that were especially influential (Barrick & Mount, 1991; Tett, Jackson, & Rothstein, 1991).

Barrick and Mount (1991) performed a meta-analysis reviewing 117 research studies that reported statistical relationships between measures of at least one of the five traits listed in the Five Factor Model (FFM) and actual job performance. To provide a more sensitised measure Barrick and Mount differentiated three forms of performance. These were, (i) job proficiency measures, (ii) training proficiency measures, and (iii) personnel data measures. Also differentiated in the study was job levels, from skilled/semi-skilled to managerial and professional. By separating different forms of job performance and job levels, Barrick and Mount provided research results that measured the relationships between personality measures and job performance equally by controlling for job level, performance, or how performance was measured. This enabled Barrick and Mount to control for additional variables that may work as moderators in the relationship between personality and job performance.

Barrick and Mount (1991) results showed that by controlling for job level, performance, and how performance was measured, personality measures tended to be good predictors of human behaviour and job performance. Amongst the traits conscientiousness emerged as the most consistent predictor of job performance/job success, with an estimated true correlation of .22 with job performance, regardless of the type of performance or job level involved. Overall people displaying high levels of this trait resulted in the perception of themselves and others of being more effective workers. Extraversion was a valid predictor also, but only across two professions managers and sales, with an estimated true correlation of .18 and .15, respectively. Extroversion was also found to correlate well with success in training .26, but only modestly with job proficiency .10 and personnel data .11. Openness to experience correlated modestly with training proficiency but not with job proficiency and personnel data.

In a more comprehensive and complex study Tett, Jackson, and Rothstein (1991) reported a meta-analysis reviewing 86 studies selected from 494 that had reported a positive relationship between measures of personality and job performance. Following on from theoretical suggestions put forward by Mischel (1968) and Bem and Allen (1974) regarding accuracy and relevance of personality traits in certain situations, the studies were separated on the basis of whether a job analysis was used. Tett et al. also analysed the studies according to four groups. (i) Recruits and incumbents, (ii) age, (iii) length of employment tenure, and (iv) civilian and military.

From the study agreeableness was found to be the strongest predictor of job performance .33, followed by openness to experience .27, Emotional Stability .22, conscientiousness, and extroversion .16. The overall estimated true correlation between personality measures and job performance was .24, higher than that reported in Barrick and Mount (1991). But more importantly those studies that used job analysis had an estimated true average correlation of .38 with job performance, as compared to .29 for those without. This shows that those traits specific to the job are more accurate predictors of job performance than traits that are not. Further, recruit performance was better predicted than incumbent performance (.30 and .20, respectively), military job performance was better predicted than civilian job

performance (.30 and .20, respectively), and neither age nor tenure mattered, nor did job level.

Although the specific results of these two meta-analysis studies (i.e. Barrick & Mount, 1991; Tett et al. 1991) were rather different, they both strongly confirmed the utility in using personality measures as predictors of job performance, and the additional improvement of these measures when coupled with a job analysis. The differences between these two studies in the true estimate correlations were thought to be due to the different jobs that were involved in the two studies, with different personality traits being more or less relevant and therefore predictive of job performance. Barrick and Mount (1993) demontrated this notion in a study of 146 civilian managers working in different U.S, Army installations. They found that when jobs were sorted in terms of their level of managerial autonomy strikingly different results were obtained than from an overall analysis. Further research in this area replicated similar findings with differences in relevant traits for various occupations (Barrick & Mount, 1996; Barrick, Mount, & Struass, 1993; Mount, Barrick, & Struass, 1994).

Within the last two decades meta-analytic research has continued to emerge supporting the relationship between personality and job performance and thus the validity of personality assessment in work related settings (e.g. Barrick, Mount, & Judge, 2001; Hogan & Holland, 2003; Judge, Bono, Erez, Locke, & Thoresen, 2002; Judge & Ilies, 2002). Furthermore these studies have shown that when used correctly i.e. when personality assessment procedures are chosen on the basis of what is needed for a job (obtained through a job analysis), they can provide incremental value to the employee selection process (Rothstein & Goffin, 2006). These findings have been assisted greatly by the development of the various factor models of personality, most notably the Five Factor Model (FFM) of Personality (McCrae & Costa, 1999).

Assessment Influenced by the Five Factor Model (FFM)

Over the last ten years much of the work conducted on personality has generally used the one model of personality over the other. This has occurred so readily during this time that there now appears to be a working consensus that the structure of personality is best based on a Five Factor Model (FFM) (McCrae & Costa, 1999). During this

time the FFM has been researched extensively and has been the focus of several annual reviews of psychological chapters (e.g. Digman, 1990), special journal issues (e.g. Costa, 1991; McCrae, 1992), and edited books (e.g. Costa & Widiger, 1993; Wiggins, 1996).

The initial development of the FFM model can be traced back to the early lexical model of William McDougall (1932), where he discussed, in the original Journal of Personality, the special meaning of "character" and "personality". In his discussion he offered an interesting conjecture: "Personality may to advantage be broadly analysed into five distinguishable but separate factors, namely, intellect, character, temperament, disposition, and temper…" Since McDougall's proposal many researchers have supplied evidence for a five-factor solution (Digman, 1990; Goldberg, 1990, 1993; John, 1990; Saucier & Goldberg, 1996). The evidence to date proposes that the FFM is a descriptive hierarchical model of the trait structure of personality, in which relatively narrow and specific traits are organized into five broad factor solutions commonly referred to as Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness (described in Table 2.1).

Table 2. 1
Description of traits referred to in the FFM.

Neuroticism / Emotional Stability: Refers to tendencies to experience negative emotions. Those high in Neuroticism generally experience anxiety, anger, and depression.
Extraversion: Refers to individual differences in the area of engagement with the external world. Extraverts enjoy being with people, are full of energy, and often experience positive emotions.
Openness / Openness to experience: Refers to a dimension of personality that distinguishes imaginative, creative people from down-to-earth, conventional people.
Agreeableness / Likability: Refers to individual differences with cooperation and harmony. Agreeable individuals value getting along with others and they are usually courteous, friendly, generous, trusting, helpful, flexible, forgiving.
Conscientiousness: Refers to the way we control, regulate, and direct our impulses. Conscientious individuals are usually careful, dependable, responsible, organized and planned.

The large influx of research on the FFM has lead to a great amount of research validating the use of the model. The research advocating the validity of the FFM can be separated into three categories.

Firstly, factor analysis research of natural language trait descriptors show good agreement across diverse cultures (Church, 2000, 2001; Goldberg, 1993; John & Srivastava, 1999; McCrae et al., 1998; McCrae & Costa, 1997). Evidence suggesting the universal nature of the FFM has led Goldberg (1990) to suggest the fundamental lexical hypothesis as evidence for the five factor structure. The fundemental lexical hypothesis suggests that " the variety of individual differences is nearly boundless, yet most of these differences are insignificant in people's daily interations with others and have remained largely unnoticed. Sir Francis Galton may have been among the first scientists to recognise explicitly the fundamental lexical hypothesis – namely that the most important individual differences in human transactions will come to be encoded as single terms in some or all of the worlds languages" (Goldberg, 1990, p. 1216).

The suggestion here is that over time humans have observed which individual differences are particularly important and thus the most important individual differences in human transactions have come to be encoded as single terms / words in some or all of the world's languages (Goldberg, 1990). The FFM captures those aspects of human transaction most significant in determining individual differences and addresses questions of how people can be expected to behave.

A second line of evidence concerns the relationship between self-ratings and ratings of an individual by others. Although the exact level of agreement varies between studies, studies have consistently shown evidence of substantial agreement between self-ratings and ratings by peers and spouses on all five factors of the FFM (McCrae & Costa, 1990).

Thirdly, there is a great deal of research that suggests the FFM is a good predictor of job performance. Moreover, the majority of studies which increased the confidence in the reliability and validity of personality measures were based on the FFM (e.g. Barrick & Mount, 1991; Tett et al. 1991). For example, Black (2000) showed evidence that Conscientiousness provided incremental validity to personnel selection methods of New Zealand Officers, Salgado (2003) showed evidence that low Conscientious employees are more likely to have deviant work behaviours and higher turnover rates.

Criticism of the FFM

In recent years research conducted on personality using the FFM has continuously risen where now the model is used readily to study areas such as stress and exercise, compulsive bullying, media use, multiple sclerosis, spinal injury, degree of intellectual engagement, the Rorschach, expatriate selection, and personnel selection (Block, 2001). The FFM's manifestation of the personality literature has led some advocates of the model to make major claims surrounding the model:

> Much of what psychologists mean by the term personality is summarized by the FFM (McCrae & Costa, 1999, p.139).

> Taken together, they (the FFM) provide a good answer to the question of personality structure (Digman, 1990, p. 436).

> The FFM has become one of the most accepted models in contemporary psychology (McCrae, 2001a, p. 108).

Due to the current trend of personality research, it is not surprising that quotations such as these have materialised. However, the claims of an emerging "consensus" about the FFM have also prompted many expressions of concern (e.g. Ben-Porath & Waller, 1992a, 1992b; Block, 1995, 2001; Collins & Gleaves, 1998; Eysenck, 1992b; Hough, 1992; McAdams, 1992; Mershon & Gorsuch, 1989; Paunonen & Jackson, 2000; Pervin, 1994; Tellegen, 1993; Waller & Ben-Porath, 1987), hence suggesting that such universal acceptance is not warranted.

With the current research and overall evidence supporting the FFM, it appears naïve to ask personality psychologists to accept this specific set of five orthogonal factors and use these factor dimensions as the conceptual structure of personality. There are three major concerns that encompass the FFM. Firstly, the issues and possibilities when using factor analysis is discussed, secondly the atheoretical nature of the FFM is discussed, and thirdly the concerns over the validity of lexical hypothesis is discussed.

The development of the FFM is heavily rooted in the method of factor analysis (Block, 1995). For some, the promise of factor analysis seems that the method would provide easy empiricism rather than rough theory to develop scientific constructs. However, when using factor analysis it is important to note the rigid logic behind the statistical method, as it can often produce mindless results (Lykken, 1971). In other words when using the method of factor analysis one must be cautious and mindful of the ways factor analysis can suggest more than is supportable. To demonstrate this a brief review of the cautionary nature needed when performing factor analysis put forward by Block (1995) is discussed.

The correlations analysed by factor analysis reflect what is commonly called communal or common variance. To the extent a variable correlates with another it is said to have communal or common variance, this means it has no unique information and is explainable by the other variable in the analysis. However, the problem with such variance is that communal or common variance observed within a particular set of variables may or may not be important in other domains. Thus, the factors summarising the communal variance may or may not be important when brought into the real world.

Extenuating this problem is that although the method of factor analysis has been used extensively for almost a century, there is still no clear, unequivocal method for deciding on how to "extract factors". There are still various rules that exist based on arguable assumptions. This is illustrated in a more general context by Cliff's (1983) remarks "There are typically an infinity of alternative sets of parameters [e.g. factor loadings] which are equally consistent with the data, many of which would lead to entirely different conclusions concerning the nature of the latent variables" (p. 122-123).

Block (1995) explains that these shortcomings highlighted about the method of factor analysis is not meant to discourage its use, as when used intelligently, it can provide valuable insight into the relationship between variables (Block, 2001). However, his contention is that proponents of FFM seem to have an over reliance on this method for advocating the use of the model in personality research. Rather than looking at other methods such as, experimental investigations, intimate and prolonged observations of

other people, neurophysiological recognitions, psychiatric insights, personal introspections, formal cognitive efforts to create a theoretical system, proponents of the FFM suggest that factor analysis is a sufficient way of deciding on a theoretical system of personality.

The second major concern regarding the use of the FFM is the need for a nomological or theoretical connection to scientific models. The FFM has been frequently criticised for being atheoretical and too reliant on empirical evidence. Recently, the FFM has been further formulated as a theory (McCrae & Costa, 1999). However, there are many personality researchers that have issues with this theory, as it can be argued that no serious approach to personality would not subscribe to these backgrounding tenets (Block, 2001). It can be further argued that the theory does not provide any information about the acceptability of the model (Eysenck, 1992). In other words, no specific theoretical consequences are entailed by the Five Factor Theory (FFT), as no sense is provided about the specific dynamics of personality. Therefore, the theory does not provide an interpretive and deductive system for personality (Block, 2001).

The third major concern that has not been settled regarding the use of the FFM is the reliance of the model on the lexical hypothesis. Previously explained in an earlier segment of this section the lexical hypothesis refers to the assumption that "the most important individual differences in human transactions will come to be encoded as single terms in some or all of the worlds languages" (Goldberg, 1990, p. 1216). However, although this is a well understood assertion with many advocates for its validity, there are some who suggest that this hypothesis should be handled with some degree of scepticism when using it for a scientific model. Goldberg himself even provides some degree of concern when using the hypothesis for the development of a model "although much can be learned about personality from language…we must not confuse the language which people use in their construction of social experience with scientific constructs that are supposed to describe, explain, and predict human behaviour…. There is reason to distrust the 'accumulated wisdom' of personality terms as a rich source… as useful starting place for scientific terminology" (Goldberg, 1982, p. 86).

The first concern regarding the credence for the lexical hypothesis, is the reliance on laypersons to specify personality descriptors. Although the fundamental assertion of the lexical hypothesis "that individual differences will be encoded in language" research in the area of cognitive development advises that there are large discrepancies between a novice and an expert in their ability to characterise and provide accurate distinctions between terms and / or variables (Peevers & Secord, 1973; Tanaka & Taylor, 1991).

This concern is perhaps accentuated by the fact that the majority of the studies empirically examining the FFM have been on undergraduates, one distinct cognitive group. Kilkowki (1976) highlighted this problem with research on the consistency of adjective usage. Kilkowski concluded that subjects did not agree about the meaning of single word trait descriptors, and that, the vocabulary skills of undergraduate populations seem untrustworthy. Further research by Goldberg and Kilkowski (1985) suggested that a novice's understanding of adjective words in our lexicon is often 'fuzzy' and definitions are needed for a novice to have full understanding. Such research suggests that personality descriptions are perhaps more likely to be valid if they are made by experts in the personality field, as they are more likely to elaborate and collaborate on definition and to provide better understanding of the words being used (Block, 1995).

The criticisms of the FFM discussed above highlight problems with this approach to the study of personality that have still not been resolved. It should be stated at this point that by no means is it suggested that the FFM is an invalid model that needs to be redeveloped and restructured for its use in personality research. Rather it is proposed that at this stage in the development of personality research the FFM constitutes one candidate, quite possibly the leading candidate, but other models still need to be researched, tested and discussed to prevent the stagnation and restriction of research conducted in the field of personality.

Assessment Influenced by Eysenck's Factor Model

Several alternative models are propounded for the study of personality, the most well-supported of which propose that there are only three fundamental dimensions of

personality (Eysenck, 1967, 1985). Eysenck's three factor model emphasized three basic dimensions of personality called – Extraversion, Neuroticism, and Psychoticism (For the remainder of this study the acronym PEN will be used to represent this model).

Extraversion and Neuroticism were first recognized by Eysenck (1947) and have since become the most well recognized factors in personality research (Pervin, 2003). These two dimensions of personality are the same dimensions proposed in the FFM (McCrae & Costa, 1992) and are dimensions that show up as important in virtually every factor-analytic trait study. Figure 1 shows the relationship between these dimensions and each dimensions associated terms proposed by Eysenck (1947).

From Figure 2.1 it is clear the the Extraversion dimension relates to differences in sociability and impulsiveness. Extraversion represents a continuous dimension where individuals can be placed, ranging from Extraversion to Introversion. The typical extravert is sociable, likes parties, has many friends, needs to have people to talk to, and does not like reading or studying by himself (Eysenck & Eysenck, 1975). They also crave excitement, take chances, often stick their neck out, act on the spur of the moment, and are generally impulsive. Extraverts tend to keep moving, may lose their temper easily, and tend not to keep their feelings under control. At the other end of the continuum a typical introvert tends to be quiet, introspective, reserved, fond of books rather than people, overall they are reserved and distant except to intimate friends. Introverts tend to plan ahead and are distrustful of impulses of the moment. They do not like excitement, are serious about aspects of everyday life, and like things to be well ordered. A typical intravert generally keeps their feelings under control, is seldomly aggressive, and do not lose their temper easily. They are reliable, slightly pessimistic, and place great value on ethical standards. The personality dimension Extraversion proposed by the model has been studied extensively and a wide variety of studies have examined the fundamental differences between extraverts and introverts. For example, these show that introverts are more easily fatigued (Eysenck, 1990), do better in school (Raine, Venables, & Mednick, 1997), are less suggestible (Acton, 2003), and are less sexually active than extraverts (Stelmack, 1990).

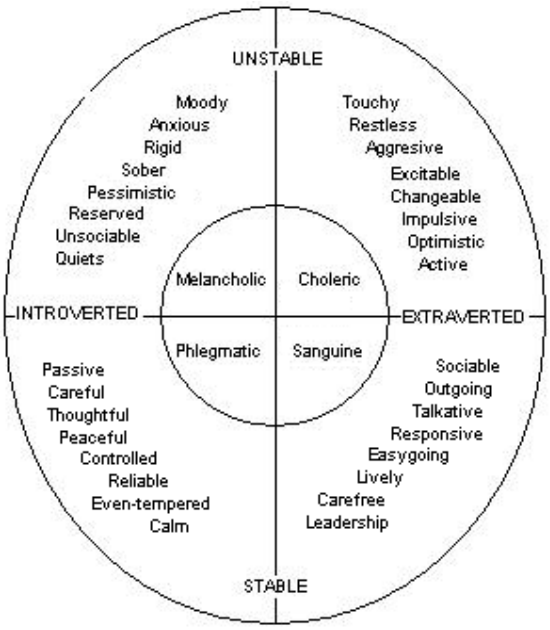

Figure 2.1
Relationship of Extraversion/Introversion and Neuroticism/Stability to earlier personality schemes.

In regards to Neuroticism, high neurotics tend to be emotionally labile, complain frequently of worry and anxiety as well as body aches (e.g. dizzy spells, headaches, stomach pain). An individual that is highly neurotic is typically moody, depressed, is likely to sleep badly, and to suffer from various psychosomatic disorders. A neurotic individual tends to be overly emotional, and finds it difficult to get back to even balanced emotions after each emotionally arousing experience. They often find it difficult to properly adjust, making them act irrationally, due to interfering emotions. When combined with Extraversion individuals are likely to be touchy and restless, to become excitable and even aggressive. At the other end of the continuum the stable individual tends to respond emotionally only slowly and generally weakly, and is likely to return to baseline levels quickly after an emotional experience. They generally are calm, even-tempered, controlled and unworried about events.

The third dimension of Eysenck's model, Psychoticism, is much more controversial than the first two discussed (Block, 1977; Bishop, 1977; Eysenck, 1985; Eysecnk & Eysenck, 1991; Kline, 1993) and as a result the exact nature of Psychoticism is less clear. However, for the most part individuals high in Psychoticism tend to be

aggressive even to loved ones, egocentric, inhumane, insensitive, hostile and unconventional. They are also likely to show a lack of empathy for others, and are often troublesome as they find it difficult to fit in. Individuals high in Psychoticism are likely to show a disregard for danger and a liking for making fools of others with the goal of upsetting them.

Eysenck's model like the FFM is based on extensive use of factor analysis on a number of existing psychometric instruments (e.g. MMPI, MPI, EPQ, EPQ-R, EPP). However, from the beginning of PEN it has had an original theory behind it on the structure of personality. Eysenck (1947) sought to locate the bases for these factors in underlying biological processes. His theory proposes a strong evolutionary significance of traits and emphasises that this translates into biological foundations for the existence of each trait proposed in PEN:

> "I feel that the major, most fundamental dimensions of personality are likely to be those on which variation has had evolutionary significance, and that this evolutionary history is likely to manifest itself in strong genetic determination of individual differences along these dimensions " (Eysenck, 1977, pp. 407-408).

The theory suggests that the major dimensions of personality are characterised by a biosocial relationship (i.e. both biological factors – hunger, thirst, sex - and social factors – laws, customs – influence human behaviour), thus, Eysenck suggests that personality has biological validity, as well as environmental validity (Eaves, Eysenck, & Martin, 1989).

Because of its theoretical backing PEN has been empirically and theoretically investigated extensively over 61 years since it was originally proposed. Although some aspects of Eysenck's theory are still questioned a large amount of research has provided support for PEN. Firstly, research indicates that additive genetic factors play a very important part in deciding a person's position across the three dimensions of PEN, in some cases contributing close to 50% of the total variance recorded on personality inventories (Eaves, Eysenck, & Martin, 1989; IvKovic et al. 2007; Matthews, Deary, & Whiteman, 2003). Secondly, research indicates that there are

differences between gender on positioning across the dimensions of PEN (Eaves, Eysenck, & Martin, 1989; IvKovic et al. 2007; Martin & Kirkcaldy, 1998; Lynn, 1991).

A third line of evidence advocating PEN is through generational and dizygotic and monozygotic twin analysis, where research indicates that Extraversion and Neuroticism can be estimated from heritability (Bouchard, 2004; Eaves, Eysenck, & Martin, 1989; Keller, Coventry, Heath, & Martin, 2005; Lake, Eaves, Maes, Heath, & Martin, 2000; Ivkovic et al. 2007). Fourthly, research suggests that the dimensions of personality proposed by PEN are good indicators of psychological disorders (Wade, Tiggemann, Heath, Abraham, & Martin, 1995). For example, there is a high genetic correlation between Neuroticism and psychological distress such as anxiety and depression (Middeldorp, Cath, Van Dyck, Boomsma, 2005). Fifthly, a large amount of research has accumulated suggesting that the personality dimensions proposed by PEN have reasonable long-term stability (Eaves, Eysenck, & Martin, 1989; Gustavsson, Weinryb, Goransson, Pedersen, & Asberg, 1997). Sixthly, research performed across cultures suggests that the dimensions of PEN are universal across cultures (Barrett, Petrides, Eysenck, & Eysenck, 1998; Martini, Mazzotti, & Setaro, 1996). Finally, research also suggests that the dimensions of personality proposed by PEN have good predictive power for work related behaviour relevent to job performance (Cox-Fuenzalida, Swickert, & Hittner, 2004; Eysenck, 1983; Eysenck & Eysenck, 1985; Martin & Kirkcaldy, 1998).

2.2 Personality Assessment in the Service of Human Resource Management Agendas

Through the introduction of comprehensive factor models and evidential research providing a linkage between personality assessment and job performance, there has been renewed enthusiasm for the use of personality assessment in industrial and military applications. Recent research indicates that personality measures, which were once used sparingly, when doubts surrounded their psychometric properties, are now commonly used throughout the world in selection batteries (McCulloch, 1993; Ryan,

McFarland, Baron, & Page, 1999; Smith & George, 1992; Taylor, Keelty & McDonnell, 2002). In addition, New Zealand organizations have been identified as being among the highest users of personality assessments (Ryan et al. 1999).

The present section briefly identifies why personality factors have been so commonly used in personnel selection especially in the Armed Forces over other personnel assessment procedures. The ethical and legal considerations which arise from this increased use of personality assessment is then discussed.

2.2.1 Personality Factors in Personnel Selection

HR specialists and psychologists view personality factors as important variables to consider when reviewing applicants for employment, as this form of testing when used on the bases of job criteria (discuss in section 2.0.2) can often add important insight to the selection process. For example, in the customer service industry individuals that score high on Emotional Stability and Agreeableness have been found to be associated with their performance in their service roles (Frei & McDaniel, 1993). However, although repeated meta-analysis research has supported the conclusion that personality predicts job performance (see Section 2.0.2), from the perspective of psychologists and HR specialists, an important question that remains is the overall utility of the assessment procedure within a selection battery. Utility analysis refers to the procedure of evaluating the value of an assessment by comparing the cost to the advantages gained from the assessment (i.e. communicating the value of an assessment). Thus, not surprisingly this form of analysis is often used to facilitate decisions on the appropriate assessment procedures to use within a selection battery.

One of the concerns surrounding the utility of any assessment is whether its predictive validity is incremental to the personnel selection techniques used. Historically this has been a sparse area of investigation, however, with studies such as Barrick and Mount (1991) recently, research on this issue has started to emerge supporting the incremental validity for personality assessment in personnel selection (Goffin, Rothstein, & Johnston, 1996; McManus & Kelly, 1999; Sackett, Gruys, & Ellingson, 1999). Of special importance results indicate that personality has incremental value

over ability tests which have historically shown to be the best predictors of job performance (Schmidt & Hunter, 1999).

The results presented above indicate that personality assessments do provide additional and relevant information on how applicants for an occupation will perform once hired for employment. However, another reason why personality measures have been increasingly used by psychologists and HR specialists is due to the comparative to costs involved in administering these types of measures. For personality assessments in self-report questionnaire format, there is no tool for the diagnosis of personality that is more economic in application. All alternative methods such as, interviews, behaviour diagnosis, and assessment centres, are cost-intensive in terms of time and resources required. Thus, for psychologists and HR specialists, personality assessment is especially attractive because of the low levels of additional resources needed.

2.2.2 Use of Personality Assessment in the Armed Forces

Personality measures are now used in a variety of industries and levels of employment. One area where personality assessment is currently used readily is the Armed forces (Army, Navy, Air force). While previous sections have examined the literature of personality assessment with particular emphasis on relationships between their utility and general job performance, the present section looks at this relationship in the setting of the Armed Forces.

Personnel selection strategies in the Armed Forces are subjected to specific and unique environmental influences that are different from traditional roles. A large portion employees within the Armed Forces are subjected to highly stressful environments and training. This requires personnel selection strategies to be effective at measuring applicants potential job performance in extreme environmental situations. Further, in addition to stressful environmental conditions, Armed Forces personnel are also subjected to frequent changes in job roles and locale, which requires considerable flexibility. Therefore, HR specialists and psychologists in the realm of the Armed Forces need to not only be concerned with the link personality

measures have with job performance but also how this link is influenced under extreme environmental conditions.

Historically the Armed Forces were one of the first industries to use personality assessment, however, in this context the link between personality and job performance has often been slightly different than in general settings. Thus, it is not surprising that researchers have spent years investigating a variety of well researched personality inventories for Armed Force selection strategies. Over the years there appears to be copious amounts of research conducted supporting the link between personality assessment and job performance in military settings (Driskell, Hogan, Salas, Hoskin, 1994; Ehud, Elishera, & Zeev, 2006; Matthews, Duckworth, & Kelly, 2007; Hartmann, Sunde, Kristensen, & Martinussen, 2003; Lung, Lee, & Shu, 2002; Medland, Green & Marder, 1965; Miller & Creager, 1960; Vickers, Hervig, & Booth, 1996; Waters, 1960). Within the settings of the Armed Forces there have been many personality inventories that have shown good levels of criterion related validity with job performance (Goyne, 2006). For example, the IPAT Nueroticism scale (NSQ), the Fleishman's Leadership Opinion Questionnaire (LOQ), all show positive correlations with outcomes in Air Traffic Controller (TAC) training in the Australian Air Force. In addition the 16PF has also shown good correlation levels with performance for Australian Air Force Pilots (Goyne, 2006).

Research findings of particular interest for New Zealand based research have been conducted using New Zealand Police recruits. These studies provide an interesting comparison with the Army officer role within New Zealand, as the New Zealand officer role is similar. In a study conducted by Black (1997) NZ Police recruits high in Conscientiousness and Extraversion were more likely to have higher levels of job performance. Further, Neuroticism was also negatively correlated with job performance. Extending on this earlier research Black (1998) assessed the difference in personality profiles between successful applicants and unsuccessful applicants for entry in the NZ Police using the NEO-PI-R. Black found that unsuccessful applicants scored higher on Neuroticism and lower on Extraversion than successful applicants.

Black's series of studies are consistent with research conducted on the overseas Armed force sample. For example, in a study conducted by Jessup and Jessup (1971)

using the Eysenck Personality Inventory (EPI) high Neuroticism and low Extraversion scores correlated well with Pilot candidates success rate in the British Royal Air Force. This finding has also been replicated in similar settings by Bartram and Dale (1982) again using the EPI. Batram and Dale's findings also indicated that results from the EPI not only correlate well with job performance but also provide incremental information.

In addition to the findings linking personality with job performance in Armed Forces settings personality dispositions have also been linked to Military personnel reactions to extreme environmental conditions (Arima, 1967; Long, Chamberlain, & Vincent, 1992; Long, 2002; Price, Monson, Callahan, & Rodriguez, 2006). Further, this relationship tends to follow similar patterns to Military job performance with high Neurotics tending to be more susceptible to difficulty in coping with extreme environments.

From the literature discussed above it is evident that because the personality measures use promises positive consequences for selection and training outcomes, they are now touted as appropriate instruments for industrial and armed forces applications. Contemporary research is substantial enough for the New Zealand Council for Educational Research (NZCER) to state in a letter to the New Zealand army "There is an abundance of empirical and theoretical support for the notion that information about personality can be usefully related to vocational decisions." However, while the positive outcomes of assessment are a key issue that psychologists and HR specialists have to be concerned with, they also have to be aware of the potential negative consequences of any assessment strategy. In addition to utility issues, one of the major issues psychologists and HR specialists must be aware of is the potential for assessment procedures to have adverse impacts.

The development of the factor models described in this chapter have helped with validation and the establishment of the utility of using objective personality assessment. With basic forms of validation supporting the use of personality assessment many researchers have now begun to investigate more complex aspects regarding the reliability, validity, and utility of personality assessment e.g. demographic variables, response bias, perception of fairness. One area of particular

importance is the influence of demographic variables on personality assessment. This is perhaps one of the most enduring topics in the history of psychometrics, as many controversies have developed in a variety of areas of psychology centering on the extent various demographically defined groups differ on important human attributes (Cronbach, 1975).

One demographic group that holds a great deal of importance in psychometrics is ethnicity/culture. Ethnic status is one of the most protected demographic groups in terms of test bias and discrimination in personnel selection, as such bias breaches many laws, morals, and ethical procedures. The current climate of personnel selection in most countries around the world including New Zealand requires psychologists and HR specialists to be aware of ethical and legal requirements of assessment strategies regarding adverse impact. The proceeding two subsections will discuss first the ethical and then the legal requirements HR specialists and psychologists are subjected to.

2.2.3 Ethical Issues

Ethical guidelines for psychologists suggests that accurate assessment of cultural differences in personality - if any - are needed to help psychologists prevent or correct "practices that are unjustly discriminatory" as required by their ethical principles. This ethical obligation is not uncommon for psychologists in other countries. However, there are reasons to propose that this ethical obligation has greater importance and significance in New Zealand.

There are two main reasons why this ethical obligation is of great importance in New Zealand. Firstly, similar to the trend of other professions the New Zealand Psychological Society firmly demonstrates its commitment to biculturalism in New Zealand. This is reflected in the recently established new Code of Ethics for all practicing psychologists working in Aotearoa/New Zealand (Bryson & Hosken, 2005). This code distinctly refers to psychologists responsibility to predict Mâori values and beliefs within New Zealand society (see, New Zealand Psychological Society, 2002). For example the code states "Psychologists…seek to be informed about the meaning and implications of the Treaty of Waitangi for their work. This

includes an understanding of the principles of protection, participation and partnership with Mâori" (section 1.3.1, New Zealand Psychological Society, 2002).

The direct reference to the Treaty in the New Zealand Psychological Society's code suggests that New Zealand psychologists should work in accordance with the Treaty and should practice in a manner which upholds these principles (Bryson & Hosken, 2005). The code suggests that New Zealand psychologists need to act in a manner that allows Mâori perspectives and preferences to be recognised, and use methods targeting Mâori specifically to reduce disparities between Mâori and NZ European. Furthermore, in dealing with disparities between Mâori and NZ European the Society's code suggests that psychologists need to act in a proactive manner when looking to reduce such disparities.

The requirements of the code need to be performed by New Zealand psychologists in the employment setting, as like many other ethnic minority groups Mâori currently experience disadvantages in work/life balance, workplace productivity, adult literacy, vocational education, and economic development. Perhaps the driving force, as well as the result of these issues, is the current unemployment rate amongst Mâori compared with NZ European in New Zealand. Recent statistics show that while the average unemployment rate in New Zealand is down to 3.6 percent, the Mâori unemployment rate remains over 8 percent (New Zealand Census, 2006). Furthermore, this statistic dramatically increases when looking at the youth unemployment rate among Mâori, which is 18 percent (New Zealand Census, 2006). With such statistics easily available to us it is easy to see why being culturally competent for psychologists in New Zealand is so important. Furthermore, because of the high use of personality assessment in employment settings throughout New Zealand (Taylor, Keelty, & McDonnell, 2002), the importance of culturally competent testing is evident.

Secondly, contemporary business is increasingly characterised by the movement of people and skill via professionally skilled emigrants and immigrants, from one economy to another (Carmon, 1996). Theoretically this is thought to help different economies continue to progress and develop, however, in many countries this results in under-employment and unemployment for emigrants and immigrants even when

they are highly skilled (International Labour Organization, 2004). New Zealand appears to follow this pattern, as New Zealand census data reveals that immigrants often have lower labour force participation than the New Zealand born population and also have a lower income (New Zealand Immigration Service, 2003). Thus, again because of this current pattern within New Zealand society it is important to perform research on selection tools (e.g. personality assessment) to establish whether these tools are potentially biased or discriminatory.

2.2.4 Legislative Issues

The accumulation of civil right legislation combined with increased motivation on the part of individuals to rectify unfair employment practices, makes the legal aspects of employment one of the most important issues in personnel selection. All aspects of the employee-employer relationship, including the initial screening, recruitment, selection, placement, compensation, training, promotion, and performance appraisals, are addressed by legislation and executive pronouncements and by legal interpretations by the courts. It is mandatory therefore for HR specialists and psychologists to use assessment strategies that fit within legislation.

New Zealand is currently signatory to a number of international laws outlawing discrimination e.g. United Nations Declaration on Human Rights; International Covenant on Civil and Political Rights; International Covenant on Economic, Social and Cultural Rights; Human Rights Act 1993, the New Zealand Bill of Rights Act 1990, the State Sector Act 1988 and the Employment Relations Act 2000. Therefore, subject to certain specific exceptions it is unlawful for New Zealand employers to discriminate on the basis of race, colour, religion, national origin, age or gender (Employment Relations Act, 2000).

Although the New Zealand Courts have struggled to express a united view on the meaning of discrimination, the following statement by the Court of Appeal in *Quilter v Attorney-General* [1998] 1 NZLR 523 has been quoted with approval in an employment context; “Discrimination generally is understood to involve differentiation by reference to a particular characteristic (classification) which characteristic does not justify the difference" (Gault, 527). This definition implies that

any assessment strategy used by HR specialists that differentiates by reference to gender, age, ethnic group or any other demographic characteristic may be subjected to legal action. Therefore, examining the issue of ethnic differences on personality assessment appears to be an important area of study, as it can provide HR specialists, psychologists and employers using various personality measures with the comfort of a thoroughly studied psychometric measure.

3. Personality and Ethnicity

The previous section showed the strong empirical evidence surrounding the reliability, validity, and utility, of personality measures which has lead to their common use in personnel selection in a variety of countries and ethnic settings (Ryan et al. 1999). However, while the criterion-related and construct validity research to date provides great support for the positive consequences for using personality assessment in personnel selection settings, it can be argued that historically certain areas of validity for personality measures have not been completely explored.

In addition to criterion-related validity when using assessment tools in employment settings HR specialists, psychologists and employers need to consider the potential for adverse impact of these tools for different groups in society. To accommodate for potential adverse impact of assessment strategies requires research to explore more complex forms of validity that focus on a broader set of equity issues for assessment administration. These issues include the fairness of test use, freedom from bias scoring and interpretation, and the appropriateness of the test based on psychological constructs or rules that underlie decision making for selection (Messick, 1989). Outlined in section 2.1 these issues are critical for psychological assessment as they also have social values and forces attached to them that are outside the measurement process when evaluative judgements and decisions are made (Messick, 1995).

Research exploring the adverse impact of assessment has focused on demographic group influences on assessment strategies. One demographic group that holds a great deal of importance in psychometrics is ethnicity/culture. The frequent use of personality measures in different ethnic settings has raised questions of whether personality measures, derived from the English language and/or Western culture, can be validly interpreted or modelled in the same way across ethnicities and cultures. Such questioning underscores the importance of understanding whether the structure and pattern of personality traits are shaped by the ethnicity/culture of an individual (e.g. educational structure, religion, and moral values). If this is so we would expect the structure of these traits to change with respect to different cultural systems, thus,

creating problems in the validity of personality measures on different ethnicities/cultures (Allik & McCrae, 2002).

The current section summarises the research performed in the area of personality and ethnicity/culture. To do this we first discuss the universality of personality traits, secondly a summary of overseas research is discussed, and thirdly a summary of New Zealand based research on this issue is discussed. Before we go further however, we need topause and consider the meaning of the term *culture*. The terms *ethnicity* and *culture* have been used inter-changeably throughout this chapter. The reason for this is because most of the research investigating personality and culture have assimilated *culture* with *ethnic* groups (Allan, 1998; Chan, 1997; Hoffman, Dana, & Bolton, 1985; Cheung & Leung, 1998; Guanzon-Lapena, Church, Carlota, and Katigbak, 1998; Knapstein, 1970; Goldberg, Sweeney, Merenda, & Hughes, 1998; Heuchert, Parker, Stumpf, & Mybrugh, 2000; Hough, 1998a; Lujan & Zapata, 1983; Ones & Anderson, 2002; Ones & Viswesvaran, 1998; Packman, Brown, Englert, Sisarich, & Bauer, 2005; Trimble, 1987; Whitworth & Perry, 1990). This is perhaps not surprising given the confusion existing about the meaning of *culture*. Richard Brislin (1983), in an Annual Review of Psychology article, noted that "like a number of concepts long studied by psychologists, such as personality, intelligence, and abnormal behaviour, there is no one definition of culture which is widely accepted" (Brislin, 1983, p. 367).

Consequently, this may explain the rationale as to why researchers have readily assimilated culture and ethnicity, as it is possible that this has occured so some form of explicit variable can provide a definition for groups representing culture. As Konstabel, Realo and Kallsmaa (2002) writes "for a study to be truly cross-cultural, researchers must at least make an attempt to specify the aspects of culture that may produce the results that are obtained or expected". Therefore the current study will continue the assimilation of the two terms ethnicity and culture, which can be understood as a group that an individual identifies with, which may have a common origin, history, destiny (Allan, 2001).

3.1 Overview of Research History

Social scientists have long debated whether Western psychology is relevant to other cultural and ethnic groups or whether such groups need to be understood using their own terms (McCrae, 2004). This has led to much scientific and public debate over the fairness of psychological assessment for minority ethnic groups since the origin of test reports in 1905 (Cronbach, 1975). For personality assessment this debate has only become persistent within the last ten years, as before this the more fundamental concepts of validity had not yet been established. Precedent to the relevance of this issue for personality focused assessments, the ill effects of psychological assessment had become most infamous with reference to Cognitive Ability Tests.

The scientific debate regarding the discrepancy of mean scores among ethnic groups on Cognitive Ability Tests goes back to the earliest stages of social science (e.g. Galton, 1869; Thorndike, 1921). Starting with the widespread use of standardised mental tests in WWI, differences among average scores for difference ethnic groups were found. Since this initial finding research has continuously demonstrated that there are significant differences between ethnic groups on Ability Tests. For example, in the United States African American and Hispanic sample groups have shown substantial differences in mean test scores compared with European Americans (Sackett, Schmitt, Ellingson & Kabin, 2001). The range of this differences is often reported to be as high as one standard deviation for the African American-European American comparison and .7 to .8 standard deviations for the Hispanic-European American comparison (Roth, Bevier, Bokor, Switzer & Tyler, 2001).

With the ill fated findings involving cognitive ability assessment it is perhaps understandable that personality research investigating the potential for adverse impact of certain ethnic groups has used a cautionary approach with minimal conclusions drawn from this line of research to date.

3.1.1 Are Personality Traits Universal?

Much of the initial work on ethnic differences in personality assessment has concentrated on establishing whether personality research administered in American

and European populations are valid and can be applied to other countries. The major problem researchers encounter when investigating this relationship is that the concepts of personality and ethnicity/culture are both pragmatic mechanisms for describing, explaining and understanding human behaviour. Moreover, personality and ethnicity/culture both attempt to answer a myriad of questions about observed, predictable, and controllable behaviour, thus creating issues when merging the two concepts together.

To help with this pragmatic issue some research focused on biosocial theories of personality have used logical arguments from evolutionary psychology. This framework of thinking suggests that, like biological "physical" traits, human behaviour and personality theory is perhaps best understood in terms of a lineage. In biology if all members of a lineage share some characteristic we may speak of the characteristic as "universal" in the lineage (Futuyma, 1979). For example, the mammal species in biology are characterized by the presence of sweat glands, hair, three middle ear bones, and a neocortex region in the brain. For biologists these shared derived characterisitics provide the basis of evolutionary taxonomies, in as much as the concordant appearance of many such traits strongly suggests they became universal in some ancestral population. Therefore, if we transfer this logic into the realm of personality we would assume that to establish universiality we need to examine whether the hierarchical structure of personality traits are consistent across ethnic groups and cultures. This is exactly what some personality psychologists have done, with much research being conducted on the hierarchical structure of personality traits in relation to models of personality (e.g. FMM, PEN).

3.1.2 Is Personality Structure Universal?

Attempting to answer this question, advocates of different types of factor models have looked to establish universality for their respective theories in the structure of personality. The methodology frequently employed has been the demonstration of high factorial structure invariance (compared with factors in the respective theories) across different ethnicities, culture, and linguistic contexts.

In the early stages of research administering this methodology inconsistent results have arisen. Commonly not recognised, McCrae and Terraciano (2005) suggest that

the major reason for this development is the different ways ethnicity and culture have been conceptualised and therefore studied. There are at least three ways research on personality may choose to conceptualise ethnicity, these are generally called ethos, national character, and aggregate personality (McCrae & Terraciano, 2005). In attempting to limit the wide range of findings the aggregated personality approach is the focus of interest in the present study. It is argued that this is the most convenient and productive approach for the assessment of personality, as standard measures of personality traits can be administered to representative samples of different ethnicities to be compared (McCrae & Terraciano, 2005)

When using the aggregated approach to personality research, there is a wealth of research in this area that supports both the personality trait structure of the FFM (e.g. Paunonen & Ashton, 1998) and the PEN (e.g. Barrett & Eysenck, 1984). For example, McCrae and Costa (1997) assessed the cross-cultural generalisability of the FFM by revising the data from six translations of the NEO PI-R. They found that German, Portuguese, Hebrew, Chinese, Korean, and Japanese samples showed similar structures after varimax rotation of five factors, suggesting the existence of the five factor structure in all six groups involved in the study. Barrett et al. (1998) also assessed the PEN across 34 different cultures. Using the Kaiser Hunka – Bianchini (PHB) procedure as a measure of factor similarities and two additional coefficients they found that, the factors in PEN are "strongly" replicable across all 34 countries used in the study.

Research such as this supports the assertion that the structure of traits hypothesised for each model, are universal across ethnic groups. Hence, suggesting that perhaps research performed in Western countries supporting the relationship between personality and job performance has utility and predictive power in other countries.

3.2 Ethnic Group Differences on Element Personality Traits

Although personality assessments are now used in a variety of countries for personnel selection, it can be argued that establishing the universality of personality traits across

ethnic groups is not satisfactory for decreeing the use of personality assessment in other countries. One important question that should be answered concerns whether personality measures used in pre-employment result in different demographic groups exhibiting different mean scores. This question has received very little attention in personality and personnel selection research, as only a small amount has been performed, mostly in the United States and Britain. The small amount of attention this area has obtained is surprising, as the quantification of group differences can facilitate our understanding of potential adverse impact on certain ethnic groups in personnel selection. Furthermore, it is perhaps "…of considerable scientific interest to discover whether various demographic groups differ in their *personalities*" (Goldberg, Sweeney, Merenda, & Hughes, 1998), as such research would likely help with further developments of personality theory and thus our understanding of human behaviour.

Perhaps one of the reasons this area has had such a low volume of interest is the problematic nature of characterising groups on ethical, conceptual, and empirical grounds. Ethically, the attribution of psychological characterisitics to different demographic groups has been the underlying reason for some of the most ugliest and violent events in history. Further, Pinker (2002) detailed in *The Blank Slate*, the possible misuse of findings advocating group differences has led social scientists to deny and ignore categorically the existence of real psychological differences among different groups. Pinker suggests that:

> "…the problem is not with the possibility that people might differ from one another…. The problem is with the line of reasoning that says that if people do turn out to be different, then discrimination, oppression, or genocide would be OK after all" (Pinker, 2002, p. 141).

However, provided that psychologists act responsibly, which means carefully qualifying their conclusions and reminding readers of the relevance of within group differences, we can reject this faulty reasoning and psychologists can ethically study possible ethnic differences in personality (McCrae & Terraciano, 2005).

3.2.1 Overseas Research

To date the research that has been performed has generated mixed results, with some studies showing distinct differences between different ethnic groups on personality scales (e.g. Barret & Eysenck, 1984; Brown, 1979; Butcher, Dahlstrom, Graham, Tellgan, & Kaemmer, 1989; Church & Lonner, 1998; Heuchert, Parker, Stumpf, & Mybrugh, 2000; Hoffman, Dana, & Bolton, 1985; Knapstein, 1970; Lujan & Zapata, 1983; McCrae, 2001, 2002; McCrae & Terraciano, 2005; Packman, Brown, Englert, Sisarich, & Bauer, 2005; Whitworth & Perry, 1990) and others indicating only trivial differences (Goldberg, Sweeney, Merenda, & Hughes, 1998; Hough, 1998a; Ones & Anderson, 2002; Ones & Viswesvaran, 1998). Although findings in this area are mixed, over the last 20 years large scale research comparing national ethnicities has arisen suggesting that there are distinctly different personality profiles between countries throughout the world.

One of the earlier and more comprehensive studies performed in this area was Barrett and Eysenck (1984) who examined the personality structure and profiles of 25 countries using the EPQ. Using non-metric multidimensional scaling Barrett and Eysenck provided data on the distance of each country from any other country based upon the combined scores of Psychoticism, Extraversion, Neuroticism and the Lie scale. For example, Barrick and Eysenck found that Australians and Chinese are most dissimilar among males, while Chinese and Ugandan are the most dissimilar among females. Hence, Barrett and Eysenck's (1984) findings suggested that different countries that have different ethnicities and cultures can be classified into different personality profiles.

In a similar study using the FFM, McCrae (2001, 2002) reported secondary analysis of data collected by other researchers from 36 cultures (defined as nation or ethnic group) using the NEO-PI-R (the content of this data were the standardised factors and facet levels means of the different cultures). McCrae separated the values into men and women subsamples and college and age subsamples. By doing this he found that again aggregated personality scores showed discrimination between cultures. In a further examination of the data McCrae analysed personality scale variances with geographical variables. This resulted in the observation that personality scales of different cultures generally had higher levels of similarity to other cultures close in

distance (i.e. neighboring countries) and higher levels of discriminality to other cultures far in distance. For example, European cultural groups showed significant differences to American cultural groups.

Extending upon McCrae (2001, 2002), McCrae and Terraciano (2005) collected data from 11,985 college students from 50 different cultural and ethnic origins on observer ratings using the 3rd-person version of the Revised NEO Personality Inventory. Like McCrae's previous studies they again found distinct differences on personality scale mean scores between cultural groups. Since these large scale studies others have followed suggesting similar findings across nations (Schmitt, Allik, McCrae, & Benet-Matinez, 2007; Schmitt, Realo, Vorack, & Allik, 2008).

The methodology of examining cross-nation differences has tended to domimate the majority of the research analysing ethnic/cultural personality differences. However, for further examination of the validity of personality assessment and personality theory these studies cannot be substituted for within country studies on ethnic group differences. Such a matter cannot be under exhorted, as without research examining within country differences the full understanding of the fairness of personality assessment used for clinical diagnosis, training, health, safety issues, and selection decisions cannot be achieved.

The number of studies focusing on within country differences on personality assessment has been significantly fewer than those examining across nations and those that have been done appear to be of minimal magnitude. However, what remains consistent with studies that have compared nations is there has been a number of studies demonstrating a difference among ethnic groups. For example, Brown (1979) compared Anglo – and Mexican American college students on the 16PF and found distinct differences between the two groups on dominance, imagination, suspiciousness, and insecurity (these lower order factors are associated with the FFM). In another study Butcher, Dahlstrom, Graham, Tellegen and Kaemner (1989) provided evidence suggesting that American Indian scores on the MMPI-2 result in significantly different personality scaling compared with White Americans. Further, these differences are relevant to 12 of the 15 personality disorder scales, with

American Indians more likely to score higher on personality disorder related scales on the MMPI-2.

In one of the few studies not focused on American or British ethnic groups, Heuchert, Parker, Stumpf and Myburgh (2000) looked at ethnic differences on personality assessment in South Africa by comparing White, Black, and Indian college students. Using Tukey post hoc they found significant differences between these groups on NEO-PI-R personality scales with the principal differences being that; Blacks scored significantly lower than Whites and Indians on Openness to Experience, Whites scored significantly higher than Blacks and Indians on Agreeableness, and Whites scored significantly higher than Blacks and Indians on Extraversion.

This trend of results suggests that when using objective personality assessment there is potential for one ethnic group within a given population to score differentially to another ethnic group to a significant degree. Generalising this trend of results to the realm of personnel selection indicates reason for concern, as this has the potential to create legislative and ethical problems and overall problems involving the effectiveness of selection strategies for an organization. However, relative to other areas of psychology where objective personality assessments are used, notably Clinical Psychology, research looking at this issue through the lens of personnel selection has been rare.

Of the few studies looking at the issue of ethnic differences on personality assessment Hough's (1998) is perhaps the most comprehensive. Although not a meta-analysis Hough (1998) provided a quantitative overview of ethnic differences on personality inventories used in personnel selection. Hough reviewed 11 inventories used for personnel selection purposes in the United States. Results from this overview reported only trivial ethnic group differences for most of the personality variables examined. However, there were a few exceptions, moderate Black – White differences on affiliation and intelligence scales were reported, Hispanic Americans scored considerably higher than Whites on response distortion scales, and Native Americans scored moderately lower than Whites on adjustment scales. Other research focusing on personnel selection have found similar results with only moderate differences

between ethnic groups found (Goldberg, Sweeney, Merenda, & Hugheds, 1998; One & Viswesran, 1998, Ones & Anderson, 2002).

The apparent lack of convergence between research examining the issue of ethnicity and personality assessment on samples relevant to personnel selection and other areas where psychomtetric assessments are conducted is encouraging for selection practitioners and employers using personality assessment for the selection program. However, this area is still drastically under investigated, as findings analysing ethnic differences on personality assessment through the lens of personnel selection are very restricted (Ones & Anderson, 2002). Furthermore, the various in-group differences on ethnicity and personality suggests that the magnitude of the differences on personality scales or traits is dependent on the groups that are compared and also the inventories used for measurement. Therefore, this suggests we have very little knowledge regarding ethnic groups that have not been studied.

3.2.2 New Zealand Based Research

Unfortunately although many personality measures are currently used in New Zealand, compared with overseas research, research addressing ethnic effects on personality assessment based on New Zealand ethnic samples has been relatively sparse. This appears to be a major shortcoming within the New Zealand literature, as it cannot be assumed that the items used for measurement of traits applying to one ethnic group will necessarily apply to another (Eaves et al., 1986). In reviewing the literature using New Zealand samples only two published studies using New Zealand samples were found (Guenole, Oleksandr, & Chernyshenko, 2005; Packman, Brown, Englert, Sisarich, & Bauer, 2005).

The focus of Guenole et al's (2005) research was to examine the suitability of the FFM for the New Zealand work context. To do this one of the steps that was incorporated was the investigation of ethnic group differences when using the model. Geunole et al. examined the personality scales of 452 employees from various occupations using Goldberg's IPIP 50-item measure (which is a direct measure of the FFM). After establishing an adequate fit of the New Zealand data to the FFM, Guenole et al. then proceeded to examine demographic variable effects. From this

analysis ethnic group differences were not found when comparing NZ European, Indian, Mâori, and Pasifika groups.

Perhaps due to the nature of Guenole et al's study (main focus was not on ethnic differences but on the suitability of the FFM model in New Zealand), Guenole et al study had a significant methodological limitation in the investigation of ethnic differences on personality assessment. The specific methodological issue is known as range restriction. This issue has had a considerable impact on research examining the interview process in selection (e.g. Roth, Van Iddekinge, Eidson, Huffcutt, & Bobko, 2002). The thinking behind this methodological issue is that a large portion of studies used samples that had already passed a previous test (employees) in the selection process resulting in those participants likely to be discriminated against already being eliminated. Thus, if employee samples are used when examining demographic influence on personality assessment it is logical that the same methodological shortcoming will arise.

In another New Zealand study Packman et al. (2005) focused on the difference in personality test scores between ethnic groups within New Zealand and between three countries. Packman et al examined personality differences between Mâori, Pacific Islander, and New Zealand European ethnic groups, and between New Zealanders, Australians, and South Africans. The study used data obtained from a consulting organization using the personality assessment 15FQ+, a popular assessment tool that examines global personality traits of an individual (i. e. FFM) and an additional 16 personality traits at the facet level. For the New Zealand sample they found differences of small to moderate magnitude at the global level for Neuroticism and Extraversion. For Neuroticism, New Zealand Europeans scored lower on average compared with both Mâori and Pacific Islanders. For Extraversion, New Zealand Europeans scored higher on average compared with Mâori and Pacific Islanders. At the facet level moderate differences were found on the Suspiciousness trait, with Pacific Islanders scoring a higher mean score than Mâori, and Mâori scoring a higher mean score than New Zealand European. They also found moderate to large differences for the Intellectance trait, with New Zealand Europeans scoring higher on average than Mâori and Pacific Islanders.

Like Guenole et al. (2005), Packman et al.'s (2005) study did have limitations. Firstly, although the two studies examined differences between personality scores among ethnic groups, they did not examine whether the differences observed translated to different employee selection outcomes for each group used in the study. Thus, they were unable to determine whether differences on personality scores lead to adverse impact for certain groups. Secondly, they did not look at the stability of the differences between ethnic groups on personality assessment scores, as only one set of data was obtained from the personality inventories used in the respective studies. These shortcomings limited the ability of both studies to determine whether the differences observed in the study were a result of personality trait differences (i.e. common score differences) among ethnic groups or differences associated with responses to the personality inventories used (i.e. measurement specific differences). This is an important shortcoming as psychological constructs such as intelligence, anxiety, and personality traits such as Extraversion have manifold and arguably value implications that may or may not be sustainable across associated measures. This is a major concern in the practice of test administration as it is important to distinguish adverse consequences of an assessment that stem from valid descriptions and invalid descriptions of individuals and groups (Messick, 1995). This shortcoming also applies to the majority of previous research conducted in this area, as most has used only one test resulting in a limited number of personality inventories being used.

Thirdly, in relation to the second limitation regarding issues of stability, both the study findings were based solely on personality inventories empirically related to the FFM, thus, making it difficult to generalise to alternative personality assessments based on alternative personality constructs (see section 2.4). This limitation for the two studies conducted in New Zealand is perhaps one of the major shortcomings of the literature on ethnic differences in personality assessment. Although there has been much research on the potential effects of ethnicity on personality most of the research looking at this issue has been conducted with the use of the FFM. As mentioned previously in section 2.4, for some this brings into question the validity and utility of such findings suggesting an ethnic impact on personality, as although the FFM has been studied extensively, it is not universally endorsed amongst psychologists (Haslam, 2007).

In addition to these limitations in New Zealand studies, there is another methodological shortcoming that perhaps relates to a large amount of research conducted on ethnic group differences and personality assessment. Most studies in this area have been inclined to focus on mean differences, the kind of relationships that historically have been of most interest to psychologists. However, there are at least two other ways that ethnicity could be associated with personality (Goldberg et al. 1998); (i) ethnicity could influence the variance of one or more personality traits. For example, New Zealand European and Mâori may show no substantial mean difference on a personality trait, but as a group New Zealand European may show more variability on the trait scores than Mâori. This type of deviation among ethnic groups is critical to trait theory, as the basic assumptions of trait theory suggest that there should be similar trait variance among ethnic groups; (ii) ethnicity might moderate the relationship between a pair of measures. For example, it could be the case that ethnicity influences the relationship between gender and personality.

Thus, not only is there a limited amount of research conducted on the relationship between personality assessment and ethnicity, there is a potential concern that present research is limited to a selected number of personality inventories used in New Zealand.

4. The Present Study

The present study addresses the issue of ethnic group differences and personality assessment using New Zealand relevant ethnic groups. Again, although it is noted that research has been conducted on this issue involving ethnic group differences on personality assessment using New Zealand participants, the shortcomings of past studies and the importance of thorough investigation into this issue cannot not be overstated. The present study provides a more systematic investigation into the issue of ethnic group differences on personality assessment than has previously been conducted in New Zealand.

The present study analyses the presence or absence of New Zealand ethnic group differences on a variety of different levels of measurement. Firstly, correlation matrices are examined to provide insight into whether the theoretical rationale and persuasive arguments linking the personality traits that are measured by personality measures are consistent across New Zealand ethnic groups. Although this form of research has been used extensively, the present research looks to examine the relationship among traits with mean differences and examine whether there is a relationship between the two levels of measurement. As previously stated this is a primary concern with respect to adverse consequences as it is important to understand whether the presence or absence of differences on assessment procedures are derived from valid or invalid measurement (Messick, 1989). In addition to this, this stage of the analysis will also allow for understanding as to how generalisable the present study findings are, as we will be able to compare these correlations with past research.

Secondly, mean trait scores differences among ethnic groups will be analysed. To do this the present study uses two personality inventories that have not previously been examined in New Zealand when analysing personality differences among ethnic groups. While exploring the mean difference among ethnic groups the present study also explores the two alternative ways that ethnicity could be associated with personality. Firstly, the variance of personality traits among New Zealand ethnic groups is analysed. Secondly, the impact of ethnicity on the relationship of age and gender on personality traits is analysed. Thus, the present design provides

examination of ethnic mean differences on personality assessment as well as the two alternative ways that ethnicity may influence personality scales.

In addition to analysing the influence of ethnicity on personality traits the present study also examines the relationship between the differences found among ethnic groups on personality assessment and personnel selection outcomes. To the authors knowledge this is the first study that has used selection data when examining the impact of ethnicity on personality. This is an imperative step in the study of ethnic differences in personality assessment, as it is the only possible method of determining whether the differences observed result in personality measures being inconsistent with the ethical and legislative obligations psychologists and HR specialists are subjected to.

To perform the present research, the study used archival data drawn from New Zealand Army Officer Selection Board (OSB). The data drawn is a group of New Zealand Army Officer applicants who completed the OSB process between 2001 and 2007. From this data information on the raw scores for personality traits measured by the EPQ-R and the GPP-I, as well as self-report information on each applicant's gender, age, ethnic group, and whether they received a job offering from the OSB selection process was gathered.

The OSB process is used to select candidates for officer training. The OSB assessors are comprised of Military Testing Officers (MTOs) and Army psychologists. The OSB process uses an assessment centre for selection of candidates for officer training. This process consists of a pre-selection phase where candidates complete a cognitive ability test, an essay-writing test, a medical screening form, and an interview with a recruiting officer. If successful at the pre-selection phase the applicants then attend a four and a half day OSB.

The personality assessment data obtained from the NZ Army OSB is administered after the pre-selection phase at the second stage of the selection process. During the interpretation of the two personality inventories, MTOs and Army psychologists use the EPQ-R trait scores to screen out applicants. For example, Psychoticism, Neuroticism, and the Lie scale is used as a "Red Flag" indicator (due to available data

these values were not available). The MTOs and Army Psychologists use the GPP-I for a different purpose as the trait scores are used to measure applicants work oriented behaviours. They are conducted over the first one and a half days that also consists of group and individual indoor and outdoor activities designed to allow assessors to assess applicants, peer ratings are gathered, and interviews conducted by the OSB. Successful applicants later attend a medical board. The results of the selection process are examined with the officer job criteria by using British Army and Australian Army officer job analysis. From this process MTOs present gradings on the selection criteria and the buddy ratings and then rank orders them. The MTOs psychologist, and deputy and president of the OSB then give gradings of In (selected) or Out (not selected).

To summarize, there are six specific questions the current study aims to address:

(i) Analyse New Zealand data to find potential relationships between personality and ethnicity. This includes comparing ethnic groups along relevant personality variables (e.g. by means of two-tailed *t* tests utilising the means as well as effect size differences and standard deviations of ethnicity defined subsets among the participants), as well as an analysis of the impact of ethnic status on gender - and age - relevant patterns in personality assessment.

(ii) Provide further insight as to whether statistical comparisons between ethnic groups reflect "real" similarities and differences.

(iii) Provide data relevant to New Zealand ethnic groups for other models besides the FFM that are currently used in personality research and personnel selection.

(iv) Check whether any relationships between personality and ethnicity can be shown to be pervasive, irrespective of choice of a (systemic) model.

(v) Examine whether differences observed on personality assessment scores significantly impact on employee selection outcomes.

(vi) Provide further data on the relationship between personality and ethnicity for New Zealand sample groups.

In the absence of a coherent body of literature on this topic and comprehensive meta-analytic studies it would be unwise to make prediction about potential findings, as there is no protection from the potential relations between personality variables and ethnicity. Therefore, the current study provides an exploratory approach.

5. Methodology

5.1 Participants

The current study used archival data drawn from New Zealand Army Officer Selection Board (OSB). The data drawn was a group of 955 New Zealand Army Officer applicants who completed the OSB process between 2001 and 2007. The database contained raw scores for personality traits, as well as self-report information on each applicant's gender, age, ethnic group, and whether they received a job offering from the OSB.

The gender breakdown of the sample was as follows: 201 females (21%), 642 males (67.2%). The ethnic group categories were American, Asian, Canadian, European, Fijian, Indian, New Zealander, NZ European, NZ Mâori, Pacific Islander, Persian, and South African. After combining the data within each of these groups, six separate groups were composed: Asian (20 or 2.1& of the sample), Pasifika (16 or 1.7% of the sample), NZ European (192 or 20.1% of the sample), Mâori (64 or 6.7% of the sample), European (96 or 10.1% of the sample), and other (25 or 2.6% of the sample). The age of applicants ranged from 16 to 58 (M= 22.66, SD= 6.63) years. Independent *t*-tests showed that there were no significant differences among ethnic groups and age, and gender and age (means given in Tables 1 and 2 respectively). All ethnic groups involved in the study had a greater amount of males (M= 69.7%) than females (M=22.4%).

Table 5.1
Mean age of different ethnic groups

Ethnic Group	Age (Mean)	SD
NZ European	23.1215	7.64836
NZ Mâori	22.5000	6.79402
Pacific Island	21.0588	3.61370
Asian	23.1000	4.61006
Total	22.9280	7.24960

Table 5.2
Mean age for males and females

Gender	Age (M)	SD
Male	22.2648	6.12111
Female	23.5628	7.78660
Total	22.5719	6.57176

5.2. Measures

The personality inventories used were the Eysenck Personality Questionnaire-Revised (EPQ-R) and the Gordon Personality Profile-Inventory (GPP-I), both are self-report personality assessments.

5.2.1. Eysenck Personality Questionnaire-Revised

The EPQ-R is a well recognised, widely used personality inventory developed to measure personality dimensions according to the Eysenck Three Factor Model (PEN). The EPQ-R consists of three scales measuring Extraversion, Neuroticism, and Psychoticism, in addition to a Lie scale designed to detect subjects who distort responses. The assessment is one of the latest questionnaires originating from the Maudsley Medical Questionnaire developed by Hans Eysenck, and is based on a large amount of research conducted using self-ratings, peer-ratings, observational studies, and analyses of psychophysiological and biochemical experiments over a forty-year period (Eysenck & Eysenck, 1991).

The Maudsley Medical Questionnaire was originally developed in 1952 to measure Neuroticism (N) in psychiatric patients. Subsequent revisions saw the development of the Maudsley Personality Inventory (MPI) and the Eysenck Personality Inventory (EPI), which added an Extraversion (E) scale to the N scale. The EPI also included a Lie (L) scale and two alternative forms to allow easier comparisons and retest (Eysenck & Eysenck, 1991). Although not included in the MMQ, MPI, and EPI, Eysenck proposed a third dimension of personality that was thought to be unmeasurable. However, this proved to be possible with the development of the

Eysenck Personality Questionnaire (EPQ), which added the Psychoticism (P) scale. The P scale was heavily criticised for its psychometric weaknesses and was therefore improved in a revision of the EPQ known as the Eysenck Personality Questionnaire – Revised (EPQ-R). Changes made to the EPQ-R largely concerned the P scale. However, two new items have been added to the E scale so there are now 23 items, one new item has been added to the N scale so there are now 24 items, and the L scale remained as it was previously at 21 items. The items on the EPQ-R are binary; requiring respondents to provide a yes-no response to each item in the assessment. The total number of items are 100. Item scores are summed to give the scale score.

EPQ-R Scales

For a more in-depth descriptive discussion regarding E, N, and P the reader can turn to pages 21-23 in this paper. However, for the sake of clarity what these scales measure will be described briefly before the more in-depth discussion of the L scale in the EPQ-R.

The 23 item E *scale* represents a measure of Extraversion that is a continuous dimension where individuals can be placed ranging from Extraversion to Introversion. The typical extravert is sociable, likes parties, has many friends, needs to have people to talk to, and does not like reading or studying by him or herself. A typical extravert craves excitement, takes chances, often sticks his/her neck out, acts on the spur of the moment, and is generally an impulsive individual. Extraverts tend to keep moving, may lose their temper easily, and tend not to keep their feelings under control. At the other end of the continuum a typical introvert tends to be quiet, introspective, reserved, fond of books rather than people, overall they are reserved and distant except to intimate friends. Introverts tend to plan ahead and are distrustful of impulses of the moment. They do not like excitement, are serious about aspects of everyday life, and like things to be well ordered. A typical introvert generally keeps their feelings under control, is seldom aggressive, and do not lose their temper easily. They are reliable, slightly pessimistic, and place great value on ethical standards.

The 24-item *N scale* measures Neuroticism. High Neurotic's tend to be emotionally labile complain frequently of worry and anxiety as well as body aches (e.g. dizzy spell, headaches, stomach pain). An individual that is highly neurotic is typically

moody, depressed, is likely to sleep badly, and to suffer from various psychosomatic disorders. A neurotic individual tends to be overly emotional, and finds it difficult to get back to even balanced emotions after each emotionally arousing experience. They often find it difficult to properly adjust, making them act irrationally, due to interfering emotions. At the other end of the continuum the stable individual tends to respond emotionally only slowly and generally weakly, and is likely to return to baseline levels quickly after emotional experience. They generally are calm, even-tempered, controlled and unworried about events.

Although there are significant changes to the 32-item Psychoticism scale in the EPQ-R it still theoretically measures the same personality construct the Psychoticism scale in the EPQ measures, Psychoticism. Because of previous uncertainty surrounding Psychoticism the exact nature of Psychoticism is less clear. However, for the most part individuals high in Psychoticism tend to be aggressive even to loved ones, egocentric, inhumane, insensitive, hostile and unconventional. They are also likely to show a lack of empathy for others, and are often troublesome as they find it difficult to fit in. Individuals high in Psychoticism are likely to show a disregard for danger and a liking for making fools of others with the goal of upsetting them.

The L scale is a measurement on the tendency of respondents who take the EPQ-R to fake good; this tendency is particularly marked when the questionnaire is administered in conditions where some tendency would deem appropriate (e.g. personnel selection). The Lie scale in the EPQ-R is constructed from items listing issues and behaviours, which are either socially desirable but infrequently practised (e.g. “Do you always wash you hands before meals?”) or frequently practised but socially undesirable (Jackson & Francis, 1999). Thus, the Lie scale in the EPQ-R permits the tendency of respondents to fake good to be diagnosed when a set of rarely performed acts are endorsed by respondents as being habitually done and when frequently performed non-desirable acts are denied by the respondents.

Factorial and experimental studies have been performed investigating the nature of L with evidence suggesting the scale has substantial factorial unity (Eysenck & Eysenck, 1970; Michaelis & Eysenck, 1971; Eysenck, Nias & Eysenck, 1971). However, in addition to measuring dissimulation the Lie scale also appears to measure

some stable personality factor. Eysenck and Eysenck (1976) suggest that if dissimulation were the only factor influencing the variance of the Lie scale, then the reliability of the scale should be a function of the size of the score obtained from the scale. In other words they suggest that when scores are low then the Lie scale should have low reliability. However, empirically this has not been found, as the reliability of the Lie scale does not appear to be significantly influenced by the score size. Despite this, recent evidence on this issue suggests that under certain conditions the score of Lie can be a reliable measure of the tendency to fake good. This condition will be discussed later when exploring the psychometric properties of the EPQ-R.

Improvements made to the Psychoticism scale: The main intention of the EPQ-R was to improve the psychometric properties of the Psychoticism scale. There were three major faults in the original P scale. These were low reliability, low range scoring, and grossly skewed distribution. Accordingly, in order to improve the Psychoticism scale in the EPQ-R, the Psychoticism scale was revised by deleting nine items and adding thirteen new items followed by factor analysis based on the original theory of Psychoticism (Eysenck & Eysenck, 1991). This results in the EPQ-R having 32 items representing the P scale instead of 25 in the original EPQ.

The new alpha coefficients for the Psychoticism scale using the new item format are in the moderate to good range with .78 for males and .76 for females, an improvement upon prior alphas of .74 for males and .68 for females in the EPQ. In addition the correlations between the new Psychoticism scale (EPQ-R) and the EPQ were calculated and the values were .88 for males and .81 for females, suggesting that findings using each test are theoretically related. The new test-retest reliabilities (one-month interval) for the Psychoticism scale using the new item format are also in the moderate to good range with .77 for males and .81 for females, which contrasts with .74 for males and .83 for females. Although these reliabilities are not as high as reliabilities achieved by Extraversion, Neuroticism, and the Lie scale, Eysenck and Eysenck (1991) argue that Psychoticism theoretically taps several different facets (hostility, cruelty, lack of empathy, etc.), that are likely to have reliabilities lower than would be true for the Extraversion, Neuroticism, and the Lie scale.

The low range scoring in the original 25 item Psychoticism scale in the EPQ was a great disadvantage with means of 3.78 ± 3.09 for males and 2.63 ± 2.36 for females resulting in a large number of zero scores on the Psychoticism scale (Kline, 1993). These results gave very little room for variation, thus weakening the Psychoticism scales discriminant validity. Using the new item format for the Psychoticism scale the EPQ-R manual reports new means of 7.19 ± 4.6 for males and 5.73 ± 3.85 for females, resulting in an improvement over the original P scale.

Finally, the skewed distribution of scores on the Psychotcism scale was a major liability in the original EPQ. The original 25-item Psychoticism scale had a skew score of 1.35 and kurtosis score of 2.65 for males (SD= 3.42) and 1.55 and 4.16, respectively, for females (SD= 2.59). This suggested large psychometric weaknesses with all values except for skewed male scores reported at more than twice the value of the standard error. Using the new item format for the Psychoticism scale, the EPQ-R manual reports a new skew score of 1.02 and kurtosis score of 1.49 for males (SD= 4.60) and .90 and 1.27 respectively for females (SD= 3.85) resulting in all values reporting results less than twice the amount of the standard error. In the EPQ-R manual Eysenck and Eysenck (1991) comment that it would be unreasonable for the Psychoticism scale to be represented with a normalised distribution, because the scale itself constitutes a departure from normality.

The Reliability and Validity of the EPQ-R

The overall reliability of the EPQ-R appears to be in the moderate to good range. The alpha coefficients are .78 (P), .90 (E), .88 (N), .82 (L) respectively for males and .76 (P), .85 (E), .85 (N), .79 (L) respectively for females. Cancelling out the Psychoticism scale these results are similar to the alpha coefficients for the original EPQ with only minor differences between individual scales in the EPQ-R and the EPQ. The EPQ-R manual also reports good test-retest reliability using a one-month interval. The reliabilities are .77 (P) .83 (E), .76 (N), .76 (N) respectively for males and .81 (P), .89 (E), .81 (N), .80 (L) respectively for females.

Information on the validity of the EPQ-R is more ambiguous than the reliability of the test, as the EPQ-R does not report a sufficient amount of detail on validity results. For example, the manual outlines only a few studies using the EPQ-R, but fails to report

any statistical evidence for most of these studies. However, the EPQ-R manual does report correlations of .88 for males and .81 for females between the EPQ original Psychoticism scale and the EPQ-R Pyshcoticism scale suggesting validity information reported for the Psychoticism scale for the EPQ-R can be drawn from the EPQ. Further, there are only minor changes to the other scales used in the test (E, N, & L), thus, EPQ research for these other scales should adequately describe EPQ-R use.

The content validity is one of the few forms of validity that the EPQ-R manual reports information on. Reported earlier when reviewing the reliability of EPQ-R, the alpha coefficients of the test fit within the moderate to good range, from .78 to .90 for males and .76 to .85 for females. These coefficients appear to be satisfactory as they are neither too high or too low, thus suggesting the items in the EPQ-R are good representatives of the test universe.

How well the EPQ-R correlates with other personality inventories is not reported in the EPQ-R manual or the EPQ manual, however, independent research has been conducted in the past suggesting good criterion validity. The technical manual for the 15FQ+ reports that the E and N scales of the EPQ correlate well with their respective 15FQ+ primary factors. The E scale correlates at a relatively good level with the 15FQ+ primary factors: Enthusiastic (.71), Socially-bold (.69), and Group-orientated/outgoing (.64). The N scale similarly correlates strongly with each related 15FQ+ primary factors: Affected by feelings (-.53), Apprehensive (.61), and Tense-driven (.56). In addition, previous research on the EPQ-R and the NEO-PI suggest that the two tests share a significant level of variance, with the majority along the dimensions of Neuroticism and Extraversion and moderate correlation between Psychoticism and Conscientiousness (Draycott & Kline, 1995). This relationship suggests that those with low scores on the Extraversion and Neuroticism scales for EPQ-R will tend to have low scores on the Neuroticism and Extraversion scales for the NEO-PI.

Good correlation levels with behavioural outcomes have also been achieved for the Extraversion, Neuroticism, and Psychoticism scales of the EPQ. Furnham (1981) conducted a study on leisure activities chosen by individuals who had been assessed using the EPQ. He found that the EPQ was a good indicator of behaviour, with

individuals identified as extraverts preferring stimulating social activities involving assertiveness and competitiveness, those identified as neurotics avoiding activities involving social interaction, and those who scored highly on the Psychoticism scale preferring situations where they could manipulate others. Further, other independent studies also provide correlations between EPQ-R scores and psychological disorders (Wade, Tiggemann, Heath, Abraham, & Martin, 1995), psychological distress (Middeldorp, Cath, Van Dyck, Boomsma, 2005), and work oriented behaviours (Eysenck & Eysenck, 1985; Martin & Kirkcaldy, 1998). Thus, there appears to be a direct relationship between responses to the test and actual observable behaviours.

Although the implications of faking are not clear some researchers argue that adding special scales to personality questionnaires that detect desirable responding can provide resistance to any potential effects of faking. The EPQ-R is an example of one such questionnaire, as it contains a Lie scale designed to detect dissimulation.

Previously stated, the Lie scale appears to measure two psychological constructs; some form of stable personality function and dissimulation of responses. Although this scale does measure the two constructs, there is considerable evidence to suggest that under certain conditions the Lie scale is a good measure of fake good (i.e. dissimulation) tendencies from respondents. Research indicates that when individuals have high motivation to fake good they inflate their Lie scale scores and suppress their Neuroticism scale scores, leading to a negative correlation between Lie and Neuroticism scale scores. This relationship has been found to hold true for children (Eysenck et al., 1965; Waters, 1968; Eysenck et al., 1971) as well as for adults (Braun & Gomez, 1966; Michaelis & Eysenck, 1971; Farley & Goh, 1976; Levin & Montag, 1987; Cowles et al., 1992). The EPQ-R manual suggests that if the Lie and Neuroticism scale correlation is significantly negative and approaches or exceeds –.5, then faking is likely to have occurred.

The EPQ-R was not specifically designed for use in the armed forces and industrial settings, rather it was designed by Eysenck and Eysenck (1991) to measure the major dimensions of personality proposed by the Eysenck Three Factor model. Although there are some studies that provide evidence that the EPQ-R has predictive power in personnel selection settings (Cox-Fuenzalida, Swickel, & Hittner, 2004), there are

only a minority of such studies. This indicates potential problems when using the EPQ-R for selection settings. However, Goeters, Timmerman, and Maschke (1993) indicate that measures such as the EPQ-R can be useful in personnel selection settings to assess how applicants react to stress, their work orientation and sociability. The New Zealand Army Officer Selection Board (OSB) use the EPQ-R for this purpose, in particular, to determine the work-oriented aspects of personality.

The EPQ-R and the EPQ manuals provides norms for a variety of occupations. However, Kline (1993) suggests that the norm samples provided are insufficiently small. To combat this problem the New Zealand Army Officer Selection Board use their own norms.

5.2.2. Gordon Personal Profile – Inventory (GPP-I)

The GPP-I is another well recognised and well used measure of personality. The GPP-I comprises of two sections which have previously been issued as separate tests. These are the Gordon Personal Profile (GPP) and the Gordon Personal Inventory (GPI). The GPP is based on a review of Cattell's (1947, 1948) personality factor research. From this review four factors were identified, Ascendancy (A), Responsibility (R), Emotional Stability (E), and Sociability (S). In addition to this the four personality scales sum to yield a measure of self-esteem (SE). Four additional factors were then developed to augment the taxonomy to provide a better basis for making predictions resulting in the development of the GPI. The four factors identified were, Cautiousness (C), Original Thinking (O), Personal Relations (P), and Vigor (V). Theoretically, the dimensions Cattell derived factorially are parallel to those of the FFM and are represented on the combined GPP-I with the FFM's Extraversion comprised into the A and S scales.

Both the GPP and GPI sections of the GPP-I contain four factor-analytically defined constructs with the four factors on the GPP merging to provide a ninth factor. Each factor construct contains forceD choice tetrads. The tetrads have one item reflecting each factor, with two of them high and two of them low preference items. Overall the GPP section contains 18 tetrads and the GPI section contains 20, with a total of 38 in the GPP-I. The two tests are comprised of 150 items.

GPP-I Scales

The *A scale* is a measure of ascendancy. The GPP-I manual describes someone who scores high on this scale as self-assured in relationships with others, generally adopting an active role in the group, and tending to make independent decisions. Low scorers on this scale are described as, those who generally play a passive role in a group, who tend to be overly dependent on others for advice, and who lack self-confidence.

The *R scale* is a measure of Responsibility. Individuals who score high on this scale generally, stick to a job assigned to them, are preserving and determined, and can be relied on. Those who score low on this scale are generally unable to stick to tasks that do not interest them, and tend to be flighty or irresponsible. The *E scale* is a measure of Emotional Stability. Individuals who score high on this scale generally are, emotionally stable, free from worries, anxieties, and nervous tensions. Those who score low on this scale are generally, excessively anxious, hypersensitive, nervous, and are frustrated easily.

The *S scale* is a measure of Sociability. Those who score high on this scale generally, like to be with and work with people, and are gregarious and sociable. Those who score low on this scale generally, have a lack of gregariousness, are restricted in social interactions, and sometimes avoid social relationships. The *SE scale* is a measure of self-esteem derived from the sum of the four scales in the GPP. This scale is primarily designed for use in counselling settings and is not used generally in the employment context. The present study does not use this scale, therefore it will not be discussed.

The *C scale* is a measure of Cautiousness. Those who score high on this scale are highly cautious, consider matters very carefully before making decisions, and do not like to take chances. Those who score low on this scale generally are impulsive, act on the spur of the moment, make hurried decisions, enjoy taking chances, and seek excitement. The *O scale* is a measure of Original Thinking. Those who score highly on this scale generally, like to work on difficult problems, are intellectually curious, and like to think about new ideas. Those who score low on this scale generally, dislike working on difficult activities, do not care particularly about acquiring knowledge, and are not interested in thought-provoking questions or discussions.

The *P scale* is a measure of Personal Relations. Those who score high on this scale generally, have faith in and trust others, are tolerant, patient, and understanding towards others. Those who score low on this scale generally, lack confidence and trust in others, tend to be critical of others, and become annoyed or irritated by others. The *V scale* is a measure of Vigor. Those who score high on this scale generally, are Vigorous and energetic, like to work and move rapidly, and are able to accomplish more than the average person. Those who score low on this scale are generally, low in energy and vitality, prefer settings of slow pace, tend to tire easily and produce low levels of work output.

The Reliability and Validity of the GPP-I

The GPP-I manual reports alpha coefficients on each scale used in the assessment. These are reported for two sample groups; a college student sample (N=134) and a male management sample (218). The coefficients reported are in the moderate to good range with . 85 (A), .87 (R), .88 (E), .86 (S), .86 (C), .87 (O), .83 (P), .85 (V) respectively for the college sample and . 82 (A), .84 (R), .82 (E), .85 (S), .83 (C), .81 (O), .83 (P), .82 (V) to .85 (S) respectively for the management sample. Using a Naval recruits sample the GPP-I manual reports test-retest (29-week interval) reliabilities in the intermediate range, the coefficients reported were .7 (A), .66 (R), .5 (E), .65 (S), .7 (C), .79 (O), .65 (P), .65 (V) respectively.

The alpha coefficients reported in the GPP-I manual also indicate that the assessment has good content validity. These coefficients suggest that the assessment provides a good representation of the test universe. The GPP-I also provides data suggesting that the assessment correlates well with other personality assessments. For example, based on a sample of 123 civil service employees the GPP-I correlated well with results from the Eysenck Personality Inventory (EPI). Theoretically expected, the most significant of these was Neuroticism negatively correlating with Emotional Stability (-.53), and Extraversion positively correlating with ascendancy (.58) and Sociability (.57). In addition, to the GPP-I correlating well with the EPI the GPP-I manual also provides empirical evidence that suggests the assessment correlates well with a variety of other measures (e.g. GZTS, SIV, WEPS, FIRO-B, Orientation Inventory, and Kuder Preference Record – Vocational).

The GPP-I manual also reports studies that indicate that scores on the GPP-I correlate with behaviour and job performance outcomes. For example, Mudge (1956) administered the GPP to 260 college students and found that the R scale on the GPP correlated significantly with students' grade point averages. In addition, the GPP-I reports a variety of studies showing the predictive validity of each scale to different performance outcomes. For example, the Original Thinking scale has been found to correlate well with scores on aptitude measures (Gordon, 1978), the Responsibility scale has been found to correlate with performance ratings by supervisors (.23) for student teachers (Warne, 1957), and the Sociability, Original Thinking, Personal Relations, and Vigor scales have all been shown to correlate with job satisfaction for special education teachers (Urban, 1972).

The GPP-I manual reports a relatively large amount of research indicating good validity in armed force settings. The manual reports a study conducted on 130 Army U.S.A officers. The study provided data indicating GPP-I scales correlated well with Academic grades (the Emotional Stability scale .18 and the Original Thinking scale .2), intercultural effectiveness (the Ascendency scale .17 and the Personal Relations scale .19), and field performance (the Ascendency scale .21, the Responsibility scale .2 and the Vigor scale .21).

The GPP-I suggests that it is not as susceptible to faking as other questionnaires that employ a yes-no response format. Studies reported in the GPP-I manual indicate only the Responsibility scale having a significantly higher mean difference in a simulated employment condition than in a simulated vocational guidance condition. In a different study, the Responsibility, Original Thinking, and Emotional Stability scales showed higher mean scores in an actual selection condition. The manual suggests that in some cases responses on some scales could be distorted, but the magnitude of the distortion is likely to be small. However, outside studies and reviews of the GPP-I suggest that the fakability of the assessment is controversial (Braun, 1963; Hess, 1998; Schwab, 1971).

Overall reviews of the GPP-I questionnaire question the practical use of the measures used to test the psychometric properties of the GPP-I, as the methods used have been

described to be out of date (Hess, 1998). Nonetheless independent reviews do suggest that the GPP-I is a good measure of personality for employment settings especially when in settings where local norms have been built (Guion, 1998; Hess, 1998). The New Zealand Army Officer Selection Board currently use their own local norms, which have been developed over a long period.

Table 5.3
List of variables used in the present study

List of Variables	What the Variable Covers	Measurement Range	Level of Measurement
Gender	Gender of Participants i.e. Male Female	Male, Female	Nominal
Ethnic group	Ethnic group subject Identifies with	NZ European, Mâori, Pasifika, Asian	Nominal
Age	Age of participants in Years	16…58	Ordinal
Job Offer	Offers by the NZ Army of employment to participant	Selected, Not Selected	Nominal
E	Participants Extraversion score on EPQ-R	0…32	Interval
N	Participants Neuroticism score on EPQ-R	0…23	Interval
P	Participants Psychoticism score on EPQ-R	0…24	Interval
L	Participants Lie score on EPQ-R	0…21	Interval
A	Participants Ascendency score on GPP-I	0...40	Interval
R	Participants Responsibility score on GPP-I	0...40	Interval
ES	Participants Emotional Stability score on GPP-I	0...40	Interval
S	Participants Sociability score on GPP-I	0...40	Interval
C	Participants Cautiousness score on GPP-I	0...40	Interval
O	Participants Original Thinking score on GPP-I	0...40	Interval
PR	Participants Personal Relations score on GPP-I	0...40	Interval
V	Participants Vigor score on GPP-I	0...40	Interval

5.3. Analysis

For each study variable (see Table 5.3), the range, mean and standard deviations were calculated. After conducting the descriptive analysis over the total data set, the same univariate statistics were also calculated separately for each relevant subgroup (e.g. females, males, NZ Mâori etc). The ethnic category groups for which data were analysed included: NZ European, Mâori, Asian, and Pasifika. Data for the 'others' category were not analysed. For each of the variables in the study at interval level of measurement (i.e. EPQ-R traits and GPP-I traits) data on the skewness and kurtosis were also computed and checked. If skewness values exceeded 1.0, log-linear transformation was performed.

The interrelationship among traits on the EPQ-R and the GPP-I were examined. In order to do this, certain trait scores were examined for bivariate relationships for any pair of scores. It is reasonable to accept international research practice that personality scale scores generated over larger numbers of test items each (10-40 items) are considered to reach interval level of measurement (in S.S. Stevens' taxonomy; Coolican, 2006). Therefore, it was safe to use Pearson product-moment correlation coefficients (*r*) to assess Bivariate relationships for any pair of personality scales.

After analysing the interrelatinship of the scales the analysis turned to the major study aims of clarifying possible similarities and differences in personality among groups defined by ethnicity. The first exploration of such differences took a variable by variable approach looking at each personality construct (EPQ-R and GPP-I scale scores) separately. The simplest way to perform such an analysis is a two-tailed *t*-test, given that none of the paired subgroups of cases are related. *t*-tests were accompanied by Levene's test, and the *d* statistic was also computed (about possible difficulties with serialised comparisons, such as *t*-tests, and an advocated more conservative approach to interpretation such as Bonferronni correction, see Section 6).

The *d* statistic in the present analysis was given added importance when examining the differences between any two groups. The *d* values were calculated by dividing the mean differences on each personality trait by the pooled standard deviation. The mean difference is expressed in terms of standard deviation units (Hunter & Schmidt, 1990). A positive *d* value is indicative of a higher mean score for the larger of the two groups being compared, while a negative *d* value is indicative of a lower mean score for the larger of the two groups. For example, a *d* of .5 means that the larger group is one-half of a standard deviation higher than the smaller group. Theoretically *d* values can range between positive and negative infinity, however, given a normal distribution, 95.44% of all effect sizes are generally found between -2.00 to +2.00 (Hunter & Schmidt, 1990). According to Cohen (1988) effect sizes of about .20 in magnitude are small, .50 in magnitude are moderate, and .80 and above are strong. In the present study effect sizes close to zero were considered to indicate equal variance between any two groups.

The reason for the decision to use *d* values are threefold. Firstly, the *d* value offers an important index to help guide decision making about selection systems. The *d* statistic provides information on the adverse impact or potential adverse impact when including information on selection ratios of selection systems (Sackett & Ellington, 1997). For example, when comparing two groups, decision makers contemplating the use of a test with a *d* value of 1.0 and a selection ratio of .25 for the majority group can expect to hire 4.7% of the minority group. Secondly, unlike *t* values the *d* statistic is independent of sample size (Roth, Bevier, Bobko, Switzer, & Tyler, 2001). *t* values and significant tests are normally affected by the sample sizes of the groups being compared in a study. This means that the detection of statistical significance between sample groups is actuated by both the magnitude of true population differences and the sample sizes used for comparison groups. Thirdly, previous studies on ethnic differences on personality scales have used effect sizes (Hough, 1998; Ones & Viswesvaran, 1998; Packman, et al., 2005; Sacket & Wilk, 1994). Therefore, using this statistic will allow the present study to compare research findings with previous research.

It is logical to proceed from a variable by variable analysis to a more holistic, multivariate analysis (Giles, 2002; Cramer, 2003) that handles personality variables collectively and looks at a holistic space of personality traits simultaneously. To answer the more holistic query about differences in personality among these groups, a multivariate analysis of variance (MANOVA) or discriminant analysis (DSC) seem to offer statistical tools (Giles, 2002; Cramer, 2003). At first sight, MANOVA may seem promising. In the current project DSC was chosen as the statistical technique.

Discriminant analysis, yielding a so-called discriminant function (Klecka, 1980; Lachenbruch, 1975), offers an appropriate multivariate technique to look at inter-group differentiation in the context of several personality variables simultaneously. When doing DSC, a larger number of variables in a study can be considered simultaneously in order to see which one(s) contribute to the discrimination between pre-defined groups, and to what extent (Klecka, 1980; Cacoullos, 1972). For the variables with the potential to discriminate groups, the procedure works through a matrix of total variances and covariances (Statsoft, 2007). On the other end, DSC relies on a matrix of pooled within-group variances and covariances. The two matrices

can then be compared via multivariate tests (such as an F test) to determine whether there are significant differences, regarding a space that includes all variables, between the groups (Klecka, 1980; Lachenbruch, 1975). In fact DSC displays mathematical similarities to MANOVA (Statsoft, 2007; Cramer, 2003).

There are several reasons to choose DSC over MANOVA (Cramer, 2003). The research is not experimental in nature. The demographic groups (in other words, the 'assignation' of each case to a certain ethnic or gender group) is dictated by the data set received; cases are certainly not randomly assigned. More importantly, MANOVA has been rarely used in projects concerning demographic correlates of personality because the status of independent and dependent variables is not self-evident. There is no assumption that variables like ethnicity, gender, or age would be correctly categorised as antecedents or precursors of personality constellations; in particular they are not considered to be "causes" of any personality pattern. DSC lends itself as a relatively more neutral but adequate technique to look at inter-group differentiation in a multidimensional data space.

Following the analysis of ethnic mean differences, multivariate analysis (MANOVA) was used to examining the interaction effects of age, gender, and ethnicity on personality trait scores. As a follow-up to the MANOVA procedure, t values from two-tailed *t*-tests, and d values were considered again. However, this time they were used separately for males and females between NZ European and Mâori. Due to data size considerations these two groups were the only ethnic groups examined in this stage of the analysis. Finally, correlation matrices, independent *t*-tests, Levene's tests and *d* statistics were computed for the job offer variable to determine what personality traits were most related to selection outcomes.

To conduct the analysis of the present study the entire analysis was performed using SPSS for windows release 11.0.1 (15 Nov 2001) version 15.

6. Results

Analysis was conducted on data from 955 officer applicants (cf. Section 5.1.). The personality inventories used in the present study were the EPQ-R and the GPP-I (cf. Section 5.2.). Before the analysis was performed univariate frequency distributions were examined for each of the variables representing personality traits. Whether trait scores were related or interrelated was examined for the total sample and whether the relationships differed for each ethnic group. To explore the differences among groups such as ethnic groups regarding personality, the first analysis used two-tailed *t* tests in inter-group comparisons. *t*-tests were accompanied by Levene's test, and effect sizes were computed. After variable by variable analysis to look at divergence between demographic (and other) groups, a multivariate analysis was performed using discriminant analysis (cf. Section 5.3).

Discriminant analysis was expected to yield some additional precision and depth beyond the output of *t*-tests and effect size scores. DSC was used to scrutinise how ethnic groups or gender groups differ from one another in the total space of personality variables. The discriminant functions resulting from this multivariate analysis may yield a more integrated picture on relative "strengths" of personality traits in "distinguishing" pairs of demographic groups. In each DSC, the total number of score level personality variables was entered collectively, to see how they contribute to the discrimination between two groups.

Gender and job offer differences were also computed on each trait score (i.e. t-test, Levene's test, effect size). Multivariate analysis was computed to analyse the interaction effects of age, ethnic status, and gender on trait scores. Finally, a statistical analysis of the relationship between ethnicity and gender as factors related to personality traits was attempted. This involved comparisons between males and females along personality traits. As mentioned in Section 5.3, all statistical techniques were executed using the statistical package SPSS (for windows release 11.0.1, 15 Nov 2001), Version 15. Numerical findings reported in this chapter are derived from SPSS output tables. The original output files are available on request.

6.1. Log-linear Transformation

Univariate frequency distributions were examined for each of the variables representing personality traits. What is presented below are measures of central tendency and measures of dispersion for each trait score (Table 6.1.1). These statistics are complemented by additional indices regarding the shape of the univariate distribution (e.g. skewness, kurtosis, minimum, maximum etc.). Table 3.0.1 shows that Psychoticism and Neuroticism on the EPQ-R and Original Thinking on the GPP-I were highly skewed. To control for this, log-linear transformation of these trait scores was computed (Table 6.1.2). The skew values were lowered for Neuroticism (-.509) and Psychoticism (-.482) on the EPQ-R. However, Original Thinking on the GPP-I increased in skewness (-9.293). As a result the further analyses performed for Original Thinking were conducted without log-linear transformation.

Log-linear transformations were also originally performed for trait scores where the kurtosis of the distribution was greater than 1.0. However, transformed scores significantly departed from normality and increased, rather than decreased, the abnormality of the scales. Thus, further analyses performed on these scales were conducted without log-linear transformation.

Table 6.1.1
Descriptive values for the EPQ-R and GPP-I scales after log-linear transformation.

	N	Minimum	Maximum	Mean	SD	Skewness		Kurtosis	
						Statistic	SE	Statistic	SE
P	955	.00	18.00	4.111	2.731	1.048	.079	1.761	.158
E	955	2.00	23.00	17.049	3.519	-.989	.079	1.024	.158
N	955	.00	19.00	4.670	3.578	1.120	.079	1.264	.158
L	955	.00	21.00	8.786	3.735	.233	.079	-.192	.158
A	955	6.00	34.00	25.354	3.718	-.515	.079	1.159	.158
R	955	11.00	36.00	28.253	3.655	-.604	.079	.613	.158
ES	955	10.00	36.00	27.156	3.621	-.439	.079	.599	.158
S	955	7.00	36.00	23.865	3.877	-.079	.079	.852	.158
C	954	5.00	36.00	23.468	4.955	-.366	.079	.236	.158
O	955	.00	40.00	30.756	4.065	-1.189	.079	4.811	.158
PR	955	1.00	40.00	26.725	4.463	-.437	.079	1.144	.158
V	955	14.00	40.00	28.823	4.189	-.393	.079	.364	.158

Table 6.1.2
Descriptive values for the P and N scales of the EPQ-R and O scale of the GPP-I after log-linear transformation.

	N	Minimum	Maximum	Mean	SD	Skewness		Kurtosis	
						Statistic	SE	Statistic	SE
P	955	.00	1.28	.643	.245	-.509	.079	.147	.158
N	955	.00	1.30	.661	.301	-.482	.079	-.181	.158
O	952	.00	1.61	1.498	.074	-9.293	.079	179.913	.158
Valid N (listwise)	952								

6.2. Interrelationships between Personality Traits

The interrelationship between traits on the EPQ-R and the GPP-I were examined. In order to do this, we took certain trait scores and examined the bivariate relationships for any pair of scores. According to Steven's level of measurement trait scores were measured at the interval level of measurement, therefore, Pearson's product-moment correlation coefficients (*r*) were chosen for this procedure. Table 6.2.1 to 6.2.4 show correlation tables of EPQ-R and GPP-I trait scores for the total number of cases in the study and separately for each ethnic group. The correlations among the four scales for the EPQ-R are presented first.

Table 6.2.1 presents the correlations among the four traits on the EPQ-R for total cases, NZ European, Mâori, Pacific Islander, and Asian groups. For the total cases in the study the correlation coefficient with the largest magnitude was -.205 and with the smallest was -.050 with an average coefficient of ±.109. Correlation coefficients representing relationships between the Psychoticism and the Lie scale and between the Neuroticism and the Lie scale show up as significant at the level of $p<.01$. However, this appears predominantly due to the large sample size in the present study with 955 participants. Both the statistically significant coefficient values report a low shared variance between these scales, with Psychoticism and the Lie scale sharing only .042 shared variance and Neuroticism and the Lie scale sharing only .034 shared variance.

Table 6.2.1
Correlation matrix for the EPQ-R scales.

	P	E	N	L
Total Cases (N = 955)				
P	-	-.005	.099	**-.205
E		-	-.112	-.050
N			-	**-.183
L				-
NZ European (N= 288)				
P	-	.030	.037	**-.207
E		-	**-.173	.008
N			-	**-.213
L				-
Mâori (N = 64)				
P	-	-.003	.216	*-.289
E		-	.001	.024
N			-	*-.260
L				-
Pasifika (N = 17)				
P	-	-.285	.010	-.193
E		-	-.073	.060
N			-	-.305
L				-
Asian (N = 20)				
P	-	-.228	.137	*-.539
E		-	**-.639	.092
N			-	-.008
L				-

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

For the NZ Europeans in the study the correlation coefficient with the largest magnitude was -.213 and with the smallest was .008, with a mean coefficient of ±.110. Table 6.2.1 displays very similar correlation coefficients to those reported for the total number of cases, with the significant correlations found suggesting only small degrees of shared variance among trait scores. Mâori on the EPQ-R had a mean correlation coefficient of ±.132, with the largest coefficient of -.289 and the smallest coefficient of .001. The Table again displays significant correlation between the Psychoticism and Lie traits and the Neuroticism and Lie traits. However, the significant correlations found suggest only a small amount of shared variance among trait scores.

For Pasifika in the study the largest correlation coefficient was -.305 and the smallest coefficient was .060, with a mean coefficient of ±.168. For Pasifika there were no significant correlations among traits. However, this may be due to the small sample size for Pasifika (N = 17), but the correlation of -.305 between the Lie scale and Neuroticism was suggestive of a negative relationship. For Asians in the present study the largest correlation coefficient was -.639 and the smallest was .082, with a mean coefficient of ±.273. Among the correlation coefficients for Asians there were two significant coefficient values. Extraversion and Neuroticism were negatively related (-.639, $p<.01$), and the Lie scale and Psychoticism were negatively related (-.539, $p<.05$). This suggests that the EPQ-R trait scores were not orthogonal for Asians.

Table 6.2.2 presents the correlations among the eight traits on the GPP-I for the total number of cases and, NZ European, Mâori, Pasifika, and Asian groups separately. For the total number of cases the correlation coefficient with the largest magnitude was -.557 and with the smallest was .008, with a mean coefficient of .205. The Table shows a large number of correlation coefficients reaching significant levels of $p<.01$ and $p<.05$. However, like the EPQ-R correlation matrix presented in Table 6.2.1 many of these significant correlations only suggested small levels of shared variance among traits (<10%). However, there were some trait scores that did suggest significant amounts of shared variance. The correlation coefficient between Emotional Stability and Sociability suggests a negative relationship (-.557, $p<.01$). The correlation coefficient for both Responsibility and Sociability are at the significance level of $p<.01$, with Responsibility having a negative relationship with Ascendency (-.494) and Sociability having a positive relationship with Ascendency (.440). The correlation coefficient between Cautiousness and Responsibility suggests a positive relationship (.447) at the significant level of $p<.01$. The correlation coefficient between Emotional Stability and Ascendency indicates a negative relationship (-.329, $p<.01$). Table 6.2.2 also reports seventeen additional correlation coefficients with levels of significance of $p<.01$ and $p<.05$. However, like the EPQ-R correlation matrix presented in Table 6.2.1 these significant correlations appear to be related to the large sample size and are not discussed as they appear to indicate only small levels of shared variance (<10%).

For the NZ Europeans in the study the correlation coefficient with the largest magnitude was -.633 and with the smallest was .008, with a mean coefficient of ±.217. The Table shows a large number of correlation coefficients reaching significant levels of p<.01 and p<.05. Like the correlation matrix for the total number of cases a large amount of these significant correlations only suggest small levels of shared variance among traits (<10%). There were some trait scores that did suggest significant amounts of shared variance. The correlation coefficient between Emotional Stability and Sociability suggests a negative relationship (-.633, p<.01). The correlation coefficient for both Responsibility and Sociability indicate relationships with Ascendency at the significance level of p<.01, with Responsibility having a negative relationship with Ascendency (-.516) and Sociability having a positive one (.478). The correlation coefficient between Cautiousness and Responsibility suggests a positive relationship (.448, p<.01). The correlation coefficient between Emotional Stability and Ascendency indicates a negative relationship (-.329) at the significant level of p<.01.

The results in Table 6.2.2 also suggest that for NZ Europeans, Ascendency had a negative relationship with Emotional Stability (-.355, p<.01) and Cautiousness (-.327, p<.01). The correlation coefficient for both Emotional Stability and Sociability with Responsibility were at the significance level of p<.01, with Responsibility having a positive relationship with Emotional Stability (.333) and Sociability having a negative one (-.335). For NZ Europeans, Sociability is also negatively related with Cautiousness (-.339, p<.01). Table 6.2.2 also reports seventeen additional correlation coefficients at the significance level of p<.01 and p<.05 for NZ Europeans. However, like the total cases correlation matrix presented in Table 6.2.1, these significant correlations appears to rise from the large sample size and are not discussed as they suggest only a low level of shared variance (<10%).

Mâori on the GPP-I had a mean correlation coefficient of ±.219, the coefficient with the largest magnitude was -.540 and with the smallest was -.004. Table 6.2.2 shows a large number of correlation coefficients for Mâori reaching significance levels of p<.01 and p<.05. Like for the total cases and NZ European correlation matrices presented in Table 6.2.2 a large amount of these significant correlations only suggest small levels of shared variance among traits (<10%). The correlation coefficient

between Emotional Stability and Sociability suggests a negative relationship (-.557, p<.01). Responsibility has a positive relationship with both Cautiousness (.521) and Vigor (.408) having coefficients at the significance level of p<.01. The correlation coefficient for both Sociability (.512) and Vigor (.412) indicate positive relationships with Ascendency at the significance level of p<.01. For Mâori, Ascendency is also negatively related to Emotional Stability (-.312, p<.05). Finally Personal Relations is positively related to Sociability (.315) and Cautiousness (.332) at the significance level of p<.05.

For Pasifika in the study the correlation coefficient with the largest magnitude was -.846 and with the smallest coefficient was -.131, the mean coefficient was ±.352. Table 3.1.2 shows high correlation coefficients among Ascendency, Emotional Stability and Responsibility. Correlation coefficients for Ascendency suggests negative relationships with both Responsibility (-.846) and Emotional Stability (-.822) at the significance level of p<.01. Further, the correlation coefficient between Emotional Stability and Responsibility suggests a positive relationship (.706, p<.01). For Pasifika the correlation coefficient between Sociability and Emotional Stability suggests a negative relationship (-.622, p<.01).

Table 6.2.2 also suggests that, for Pasifika, Cautiousness is positively related to Responsibility (.569, p<.05). The correlation coefficient for Sociability is positively related to Ascendency (.502, p<.05). Finally both the correlation coefficients between Cautiousness and Personal Relations (.525) and Original Thinking and Vigor (.539) suggest positive relationships at the significance level of p<.05.

For Asians in the study the largest correlation coefficient was -.739 and the smallest coefficient was .017, with a mean coefficient of ± .276. For Asians in the study Ascendency has a positive relationship with Sociability (.739) and Vigor (.614) with coefficients at the significance level of p<.01. Vigor also has a positive relationship with Original Thinking (.662, p<.01).

Table 6.2.2
Correlation matrix for the GPP-I traits.

	A	R	ES	S	C	O	PR	V
Total cases (N = 955)								
A	-	-.494**	-.329**	.440**	-.294**	.253**	-.074*	.216**
R		-	.279**	-.291**	.447**	.008	.131**	.263**
ES			-	-.557**	.189**	-.018	.203**	-.058
S				-	-.266**	.134**	.036	.121**
C					-	-.021	.218**	-.016
O						-	.042	.166**
PR							-	-.178**
V								-
NZ European (N = 288)								
A	-	-.516**	-.355**	.478**	-.327**	.154**	-.082	.263**
R		-	.333**	-.335**	.448**	.082	.130*	.204**
ES			-	-.633**	.264**	.058	.166**	-.042
S				-	-.339**	.071	.013	.156**
C					-	.061	.212**	-.059
O						-	.048	.116*
PR							-	-.122*
V								-
Mâori (N = 64)								
A	-	-.293*	-.302*	.512**	-.179	.279*	.067	.412**
R		-	.170	-.122	.521**	.022	.167	.408**
ES			-	-.540**	.194	.188	.202	-.083
S				-	-.145	.120	.315*	.100
C					-	-.004	.332**	.111
O						-	.045	.244
PR							-	-.067
V								-
Pasifika (N = 17)								
A	-	-.846**	-.822**	.502*	-.453	.197	-.159	.186
R		-	.706**	-.423	.569*	.155	.154	-.167
ES			-	-.622**	.351	-.239	.204	-.131
S				-	-.385	.284	-.124	.174
C					-	-.219	.525*	-.296
O						-	-.263	.539*
PR							-	-.150
V								-
Asian (N = 20)								
A	-	-.307	-.442	.739**	.230	.535*	-.098	.614**
R		-	.548*	-.220	.216	.075	.107	.341
ES			-	-.375	-.052	-.028	.058	.017
S				-	.063	.518*	.253	.552*
C					-	.161	.118	.092
O						-	-.285	.662**
PR							-	.018
V								-

** Correlation is significant at the 0.01 level (2-tailed).
is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Table 6.2.3 presents the correlations among traits on the EPQ-R and on the GPP-I for total cases, NZ European, Mâori, Pasifika, and Asian groups. For total cases the

largest correlation coefficient was .495 and the smallest coefficient was .007 with a mean coefficient of .169. Like in the previous two correlation matrices listed in Table 3.1.1 and 3.1.2 the correlation coefficients for the total number of cases shows a large number reaching significant levels of p<.01 and p<.05, however, a large amount of these significant correlations suggest only small levels of shared variance between scales (<10%). There are four correlation coefficients that suggest more than 10% shared variance. Table 6.2.3 shows that Extraversion had a positive relationship with Sociability (.495) and Ascendency (.480) at the significance level of p<.01. Extraversion for the total number of cases also had a negative relationship with Cautiousness (-.312, p<.01). The correlation coefficient between Psychoticism and Cautiousness had a negative relationship (-.329, p<.01).

For NZ Europeans the largest correlation coefficient was .582 and the smallest coefficient was .000 with a mean coefficient of .167. Table 6.2.3 reports significant correlations between Extraversion and Sociability (.582) and Extraversion and Ascendency (.534), both positive relationships at the significance level of p<.01. Extraversion for the total number of cases also had a negative relationship with Cautiousness (-.361, p<.01) and a positive relationship with Vigor (.343, p<.01) and Emotional Stability (.321, p<.01). The correlation coefficient between Psychoticism and Cautiousness had a negative relationship (-.342, p<.01).

For Mâori in the study, the largest correlation coefficient was .517 and the smallest coefficient was -.030, with a mean coefficient of ± .209. For Mâori the correlations coefficient between Extraversion and Sociability suggest a positive relationship (.582, p<.01). Psychoticism had a negative relationship with both Responsibility (-.423) and Cautiousness (-.375) with both correlation coefficients at the significance level of p<.01. The correlation coefficients between the Lie scale and Cautiousness (.401) and the Lie scale and Responsibility (.341) indicate positive relationships at the significance level of p<.01. Neuroticism had a negative relationship with both Responsibility (-.311), Personal Relations (-.356), and Vigor (-.380) with correlation coefficients at the significance level of p<.01. Finally, the correlations coefficient between Extraversion and Ascendency suggest a positive relationship (.381, p<.01).

Table 6.2.3
Correlation matrix among the EPQ-R traits and the GPP-I traits.

	A	R	ES	S	C	O	PR	V
Study sample (N = 955)								
P	.058	-.246**	-.040	.007	-.329**	-.082*	-.164**	-.132**
E	.480**	-.197**	-.290**	.495**	-.312**	.176**	.029	.289**
N	-.154**	-.099**	-.272**	.032	-.074*	-.102**	-.207**	-.229**
L	-.116**	.227**	.042	-.053	.222**	-.021	.144**	.071*
NZ European sample (N = 288)								
P	.068	-.232**	-.050	-.015	-.342**	-.143*	-.190**	-.079
E	.534**	-.216**	-.321**	.582**	-.361**	.101	.070	.343**
N	-.202**	-.052	-.277**	.017	-.043	-.062	-.171**	-.239**
L	.000	.179**	.050	.013	.133*	.069	.096	.088
NZ Mâori sample (N=64)								
P	.174	-.423**	-.127	-.030	-.375**	-.122	-.159	-.164
E	.381**	-.111	-.307*	.517**	-.073	.137	.085	.202
N	-.121	-.311*	-.270*	-.043	-.217	-.172	-.356**	-.380**
L	-.031	.341**	.181	-.035	.401**	.087	.145	.221
Pacific Islander sample (N = 17)								
P	.116	-.480	.068	-.097	-.317	-.324	.148	.172
E	.451	-.378	-.518*	.622**	-.470	.171	-.239	.397
N	-.185	.203	.264	-.313	.126	-.003	-.131	.058
L	-.536*	.534*	.382	-.204	.438	-.189	.454	.061
Asian sample (N = 20)								
P	.068	-.258	-.194	-.113	-.394	-.009	-.395	-.134
E	.539*	-.070	-.199	.479*	-.084	.469*	-.027	.574**
N	-.541*	.007	.234	-.584**	.130	-.506*	-.150	-.623**
L	-.158	.063	-.140	-.101	.329	-.053	.458*	-.127

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

For Pasifika in the study the largest correlation coefficient was .622 and the smallest coefficient was -.003, with a mean coefficient of ± .283. For Pasifika the correlation coefficient between Extraversion and Sociability suggest a positive relationship (.622, $p<.01$). The Lie scale had a negative relationship with Responsibility (-.536, $p<.05$), the Lie scale also had a positive relationship with Responsibility (.534, $p<.05$). Finally, the correlations coefficient between Extraversion and Sociability suggest a negative relationship (-.518, $p<.05$).

Table 6.3.1
T-test and d statistic values of scale scores for NZ European and Mâori.

	Levene's Test					
	F	Sig.	t	df	Sig. 2-tailed	d Statistic
EPQ-R Scales						
P	0.479	0.489	-1.254	350	0.211	-0.173
			-1.232	91.374	0.221	
E	1.853	0.174	-0.626	350	0.532	-0.086
			-0.657	98.477	0.513	
N	0.229	0.633	-0.333	350	0.74	-0.046
			-0.325	90.94	0.746	
L	1.547	0.214	-0.897	350	0.37	-0.124
			-0.824	85.689	0.412	
GPP-I Scales						
A	0.121	0.728	0.481	350	0.631	0.066
			0.5	97.291	0.618	
R	1.15	0.284	1.172	350	0.242	0.162
			1.17	92.899	0.245	
ES	0	0.989	-0.749	350	0.454	-0.103
			-0.763	95.024	0.447	
S	0.175	0.676	0.622	350	0.534	0.086
			0.609	91.041	0.544	
C	0.408	0.524	-0.666	350	0.506	-0.092
			-0.665	92.984	0.508	
O	1.939	0.165	2.116	350	0.035*	0.292
			1.828	81.476	0.071	
PR	0	0.983	-0.528	350	0.598	-0.073
			-0.526	92.719	0.6	
V	2.734	0.099	1.211	350	0.227	0.167
			1.101	84.976	0.274	

* Correlation is significant at the 0.05 level (2-tailed).

6.3. Mean Ethnic Differences on Personality Traits

6.3. 1. Ethnic Differences on Single Personality Traits

Our aim (cf. Chapter 4) was to analyse data in order to locate any differences in personality traits indicating a disparity between two ethnic groups. The analysis proceeded in several stages. First, inter-group comparisons with a personality trait as a single variable were performed by using a two-tailed *t*-test, Levene's test, and the *d* statistic. The variables examined for the comparisons were the EPQ-R and GPP-I trait scores; these variables all reach the interval level of measurement. Tables 6.3.1 to 3.2.6 present group comparisons on each personality scale and display *t*-test, effect size, and the F ratio from Levene's test. The most important differences were

considered to be differences that were significant or exceeding 0.2 standard deviations (SD). In addition, Levene's F ratio reaching $p<.05$ were also deemed important, however, there were no significant SD differences found in this stage of the analysis from the Levene's F ratio.

When making multiple hypothesis tests like the ones reported in these results, in order to control for the likelihood of obtaining statistically significant results by chance, it is common to compute a Bonferroni significance level (Glass & Hopkins, 1996). The Bonferroni is a statistical adjustment for the multiple comparisons that are made in a study where there are a large number of investigations computed simultaneously on the same data set. It is designed to raise the standards of proof needed for each of the significant results and thus, lowers the chance of Type I errors in an analysis (Hinkle, Wiersma, & Jurs, 2003).

The Bonferroni correction procedure is simply a procedure where the alpha level of a significant test is divided by the amount of outcomes in a test (Agresti & Finlay, 1997). For example, for the present study, on the GPP-I we measured the impact of ethnicity on eight personality traits. Instead of testing the traditional .05 alpha level, with Bonferroni corrections we would test at alpha = .05/8 = .0625. This would ensure that the overall chance of making a Type I error is still less than .05.

The analysis performed in the present study did not compute a Bonforroni correction. The decision not to apply this correction on the data was for two reasons; (1) the nature of this study is exploratory, therefore, it is possible that the implementation of this procedure could cause a substantial loss in the precision of the research findings, (2) when the sample size is small and the number of tests is large, the power of multiple testing procedures is extremely poor, this can lead to only unusually strong effects being observed while missing an uncontrollable number of significant findings. The present study had a small amount of participants representing Mâori, Pasifika, and Asian groups. Therefore, performing the Bonferroni correction had the possibility of making the risk of Type II errors even higher (Nakawaga, 2004). As a result the results report statistics that have not had Bonferroni correction computed.

Between NZ Europeans and Mâori there was only one significant mean difference. This was for Original Thinking on the GPP-I. This trait reports a significant difference at $t(350) = 2.12$, $p < .05$. The trait had an effect size of $d = .292$, suggesting that on this scale NZ Europeans tend to score higher than Mâori (see Figure, 6.3.9). There was no significant differences between NZ European and Mâori in terms of SD from Levene's analysis.

Between NZ European and Pasifika ethnic groups there were two traits with significant mean differences. Trait scores for Personal Relations was significantly different between the two group $t(303) = -2.35$, $p<.05$, with Pasifika scoring higher than NZ Europeans ($d = -.585$, see Figure 6.3.11). Trait scores for Vigor were also significantly different between the two groups $t(303) = 1.99$, $p<.05$, with NZ Europeans scoring higher than Pasifika ($d = .497$, see Figure 6.3.12).

Table 6.3.2
T-test and d statistic values of scale scores for NZ European and Pasifika.

	Levene's Test					
	F	Sig.	t	df	Sig. 2-tailed	d Statistic
EPQ-R Scales						
P	0.029	0.865	-1.069	303.000	0.286	-0.267
			-1.150	18.298	0.265	
E	0.181	0.671	0.444	303.000	0.657	0.111
			0.471	18.221	0.644	
N	2.108	0.148	-0.632	303.000	0.528	-0.158
			-0.739	18.793	0.469	
L	0.033	0.855	-0.864	303.000	0.388	-0.216
			-0.821	17.726	0.422	
GPP-I Scales						
A	0.016	0.900	0.740	303.000	0.460	0.185
			0.733	17.902	0.473	
R	0.168	0.682	-0.979	303.000	0.328	-0.244
			-0.953	17.823	0.354	
ES	0.013	0.909	-1.204	303.000	0.230	-0.301
			-1.252	18.123	0.227	
S	0.921	0.338	0.993	303.000	0.322	0.248
			0.804	17.196	0.432	
C	0.133	0.716	0.221	303.000	0.825	0.055
			0.197	17.493	0.846	
O	0.848	0.358	0.086	303.000	0.931	0.022
			0.109	19.361	0.914	
PR	0.097	0.756	-2.345	303.000	0.020*	-0.585
			-2.429	18.106	0.026	
V	3.846	0.051	1.991	303.000	0.047*	0.497
			2.898	20.721	0.009	

* Correlation is significant at the 0.05 level (2-tailed).

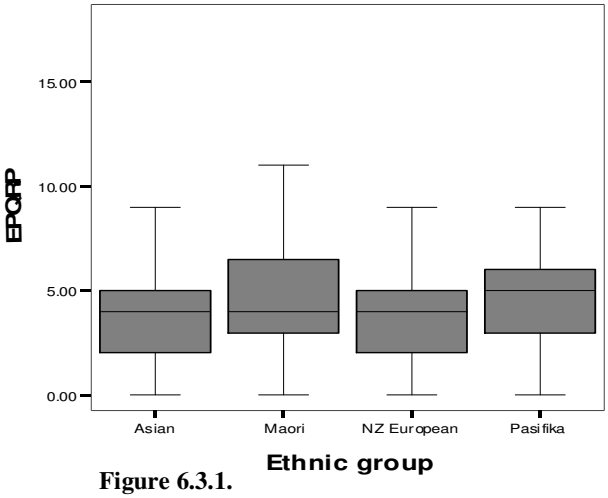

Figure 6.3.1.
Comparison among Ethnic groups on Median and SD values for Psychoticism on EPQ-R.

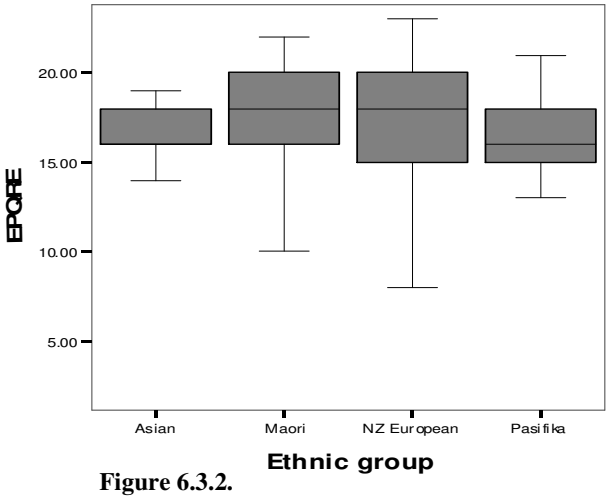

Figure 6.3.2.
Comparison among Ethnic groups on Median and SD values for Extraversion on EPQ-R.

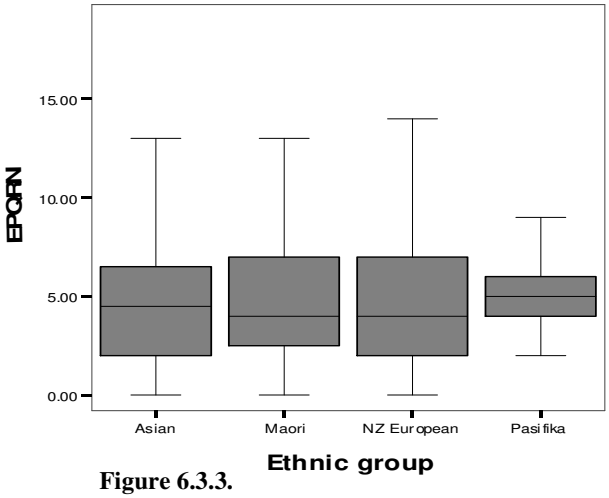

Figure 6.3.3.
Comparison among Ethnic groups on Median and SD values for Neuroticism on EPQ-R.

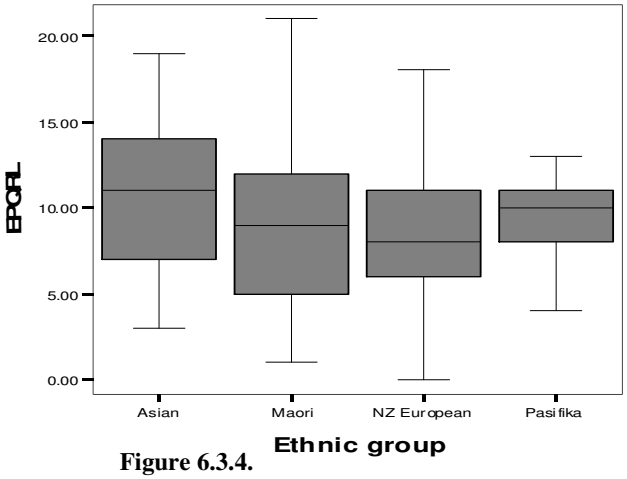

Figure 6.3.4.
Comparison among Ethnic groups on Median and SD values for the Lie scale on the EPQ-R.

In addition to the two significant differences reported between NZ Europeans and Pasifika the effect sizes indicated some additional differences. The effect size for Psychoticism on the EPQ-R (d = -.267), indicates that Pasifika tend to score higher for this trait than NZ Europeans (see Figure 6.3.1). The effect size for the Lie scale on the EPQ-R (d = -.216), indicates that Pasifika tend to score higher on this scale than NZ Europeans (see Figure 6.3.4).

For Responsibility on the GPP-I the effect size is d = -.244 suggesting that Pasifika tend to score higher on this trait than NZ Europeans. The effect size for Emotional Stability (d = -.301), indicates that Pasifika tend to score higher on this scale than NZ Europeans. Finally, the effect size for Sociability on the GPP-I (d = .248) indicates that NZ Europeans tend to score higher on this trait than Pasifika (see Figure 6.3.8).

Between NZ European and Asian ethnic groups there was one trait with a significant difference between groups. The Lie scale on the EPQ-R had significant differences between ethnic groups $t(306)= -2.511$, $p<.05$, with Asians scoring higher than NZ Europeans (d = -.580, see Figure 6.3.4). In addition to the significant difference reported between NZ Europeans and Asians the effect sizes indicated some additional differences. The effect size for Neuroticism on the EPQ-R (-.201) indicates that Asians tended to score higher on this trait than NZ Europeans (see Figure, 6.3.2). Ascendency on the GPP-I had an effect size of d = .267, indicating that NZ Europeans

tended to score higher on this trait than Asians (see Figure 6.3.5). The effect size for Sociability scale (*d* = .206) indicated that NZ Europeans tended to score higher on this trait than Asians (see Figure 6.3.8). Finally, the effect size for Vigor on the GPP-I (*d* = .391) indicated that NZ Europeans tended to score higher on this trait than Asians (see Figure 6.3.12).

Between Mâori and Pasifika ethnic groups there was one trait with a significant mean difference at p<.10. The Personal Relations scale on the GPP-I reports a significant difference between ethnic groups at *t*(79) = -1.878, p<.10, indicating that Pasifika tended to score higher on this trait than Mâori (*d* = -.513, see Figure 6.3.11). In addition to the two significant differences reported between Mâori and Pasifika the effect sizes indicated some additional differences. The effect size for Emotional Stability on the GPP-I (*d* = -.203) indicated that Pasifika tended to score higher on this trait than Mâori (see Figure 6.3.7) The effect size for Original Thinking on the GPP-I (*d* = -.244) indicated that on this scale Pasifika tended to score higher on this trait than Mâori (see Figure 6.3.9).

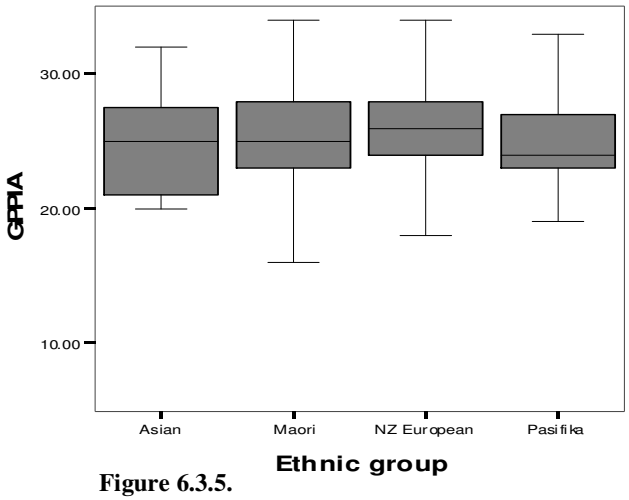

Figure 6.3.5.
Comparison among Ethnic groups on Median and SD values for Ascendancy on GPP-I.

Table 6.3.3

T-test and d statistic values of scale scores for NZ European and Asian.

	Levene's Test		t-test for Equality of Means			
	F	Sig.	t	df	Sig. 2-tailed	d Statistic
EPQ-R Scales						
P	0.21	0.647	-0.121	306	0.904	-0.028
			-0.109	21.142	0.914	
E	2.79	0.096	0.567	306	0.571	0.131
			0.693	23.394	0.495	
N	0.989	0.321	-0.869	306	0.385	-0.201
			-0.71	20.69	0.486	
L	1.201	0.274	-2.511	306	0.013*	-0.580
			-2.187	20.966	0.04	
GPP-I Scales						
A	2.045	0.154	1.157	306	0.248	0.267
			0.877	20.414	0.391	
R	0.732	0.393	-0.001	306	0.999	0.000
			-0.001	21.453	0.999	
ES	0.106	0.744	-0.351	306	0.726	-0.081
			-0.342	21.552	0.736	
S	0.097	0.756	0.889	306	0.375	0.206
			0.777	20.98	0.446	
C	0.676	0.412	-0.553	306	0.581	-0.128
			-0.466	20.816	0.646	
O	0.699	0.404	0.495	306	0.621	0.114
			0.437	21.032	0.667	
PR	0.731	0.393	-0.042	306	0.966	-0.010
			-0.048	22.762	0.962	
V	0.018	0.893	1.69	306	0.092	0.391
			1.65	21.573	0.113	

* Correlation is significant at the 0.05 level (2-tailed).

Between Mâori and Pasifika the effect sizes for Vigor on the GPP-I (*d* = .295) indicated that Mâori tended to score higher on this trait than Pasifika. Finally, the effect size for Responsibility on the GPP-I (*d* = -.403), indicated that Pasifika tended to score higher on this trait than Mâori (see Figure 6.3.6).

Between Mâori and Asian ethnic groups there were no traits with significant mean differences. However, effect sizes on Table 6.3.4 indicated some difference. The effect size for Extraversion on the EPQ-R (*d* = .240) indicated that Mâori tended to score higher on this trait than Asians (6.3.2). The effect size for the Lie scale on the EPQ-R (*d* = -.400) indicated that on this scale Asians tended to score higher on this trait than Mâori (see Figure 6.3.4).

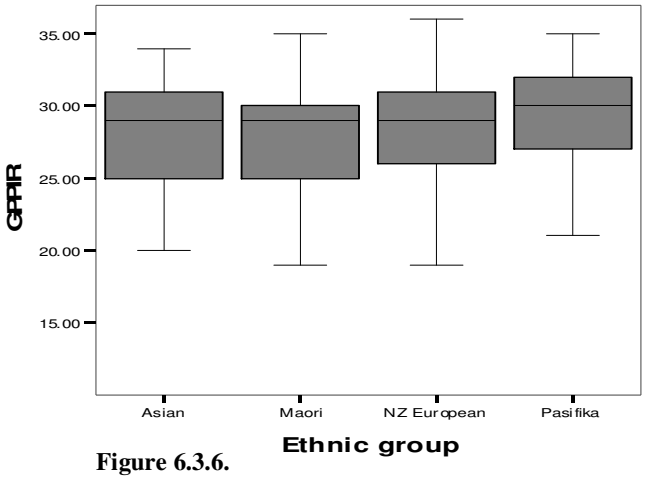

Figure 6.3.6.
Comparison among Ethnic groups on Median and SD values for Responsibility on GPP-I.

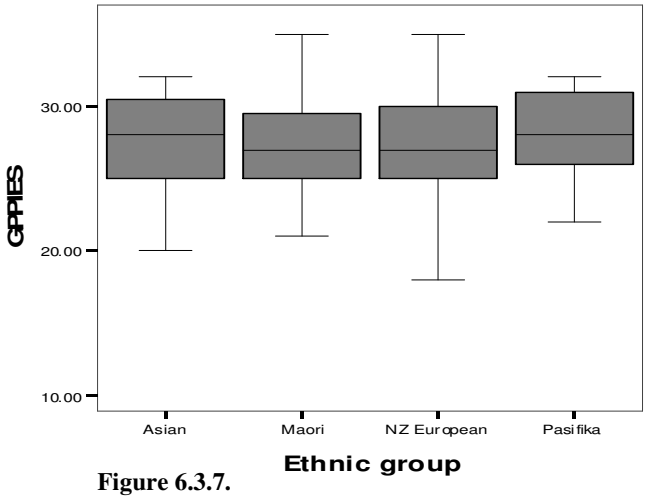

Figure 6.3.7.
Comparison among Ethnic groups on Median and SD values for Emotional Stability on GPP-I.

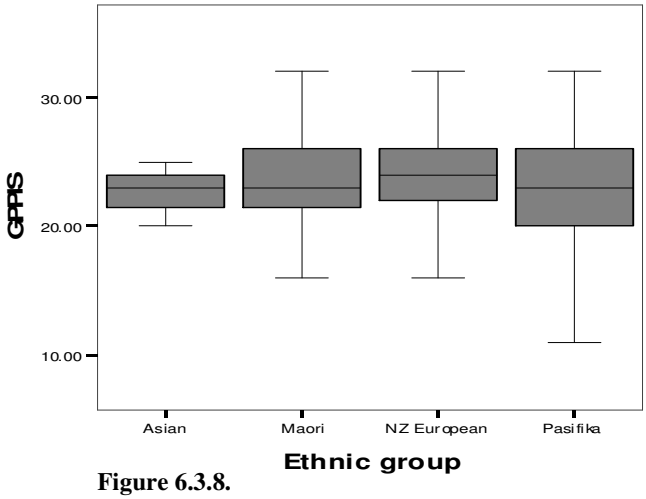

Figure 6.3.8.
Comparison among Ethnic groups on Median and SD values for Sociability on GPP-I.

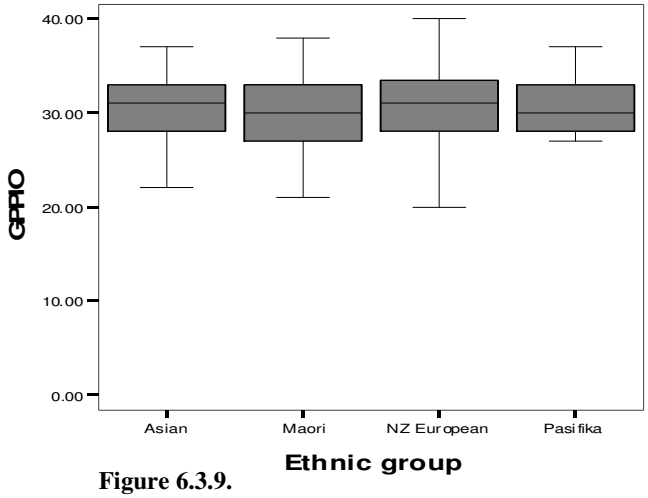

Figure 6.3.9.
Comparison among Ethnic groups on Median and SD values for Original Thinking on GPP-I.

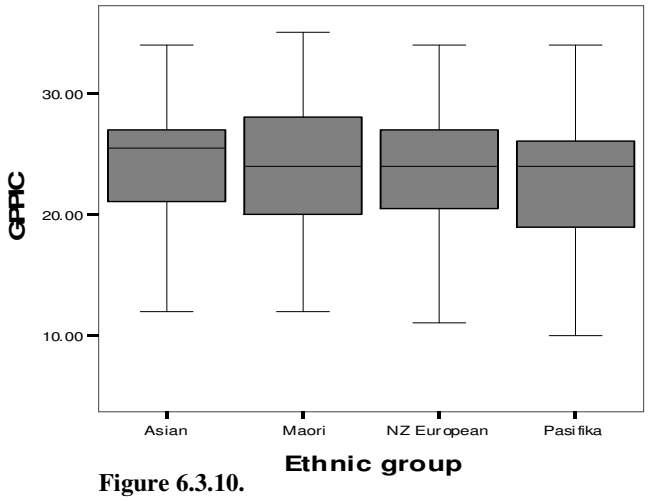

Figure 6.3.10.
Comparison among Ethnic groups on Median and SD values for Cautiousness on GPP-I.

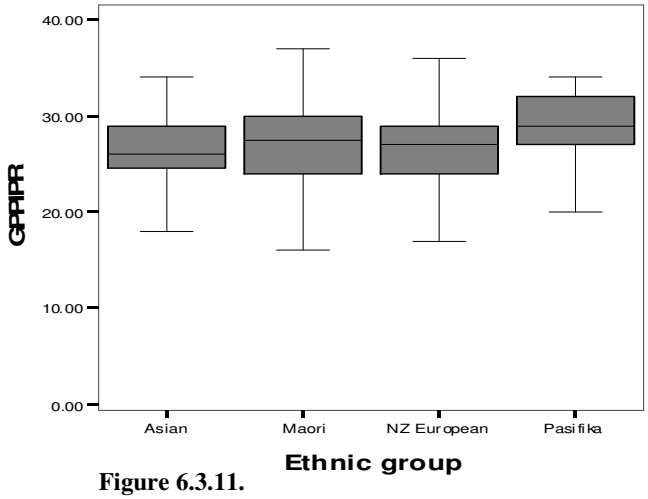

Figure 6.3.11.
Comparison among Ethnic groups on Median and SD values for Personal Relations on GPP-I.

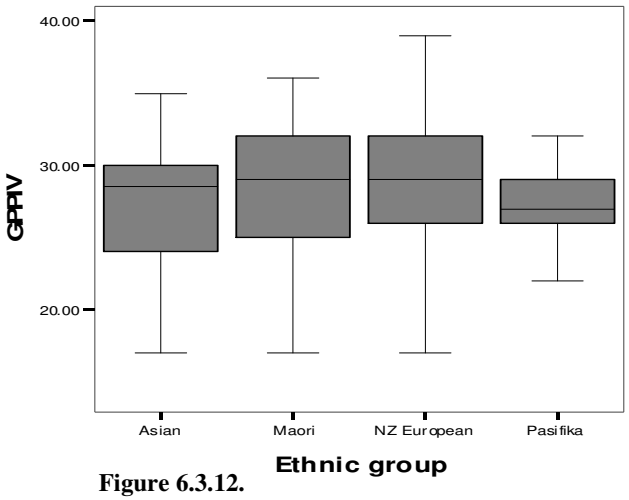

Figure 6.3.12.
Comparison among Ethnic groups on Median and SD values for Vigor on GPP-I.

Table 6.3.4
T-test and d statistic values of scale scores for Mâori and Pasifika.

	Levene's Test		t-test for Equality of Means			
	F	Sig.	t	df	Sig. 2-tailed	d Statistic
EPQ-R Scales						
P	0.298	0.587	-0.333	79	0.74	-0.091
			-0.355	27.575	0.725	
E	0.092	0.763	0.773	79	0.442	0.197
			0.768	24.95	0.45	
N	2.486	0.119	-0.404	79	0.687	-0.110
			-0.457	30.339	0.651	
L	0.499	0.482	-0.29	79	0.773	-0.079
			-0.303	26.795	0.764	
GPP-I Scales						
A	0.101	0.752	0.457	79	0.649	0.125
			0.438	23.896	0.665	
R	0.035	0.852	-1.478	79	0.143	-0.403
			-1.454	24.664	0.158	
ES	0.012	0.914	-0.745	79	0.458	-0.203
			-0.751	25.46	0.459	
S	0.405	0.527	0.558	79	0.579	0.152
			0.495	21.983	0.625	
C	0.001	0.982	0.525	79	0.601	0.143
			0.488	23.032	0.63	
O	1.571	0.214	-0.891	79	0.376	-0.244
			-1.174	41.776	0.247	
PR	0.074	0.786	-1.878	79	0.064	-0.513
			-1.93	26.109	0.065	
V	6.368	0.014	1.078	79	0.284	0.295
			1.476	45.965	0.147	

Between Pasifika and Asian ethnic groups there was one trait with a significant mean differences at p<.10. Personal Relations on the GPP-I had significant differences between ethnic groups $t(37) = 1.923$, p<.10, with Pasifika scoring higher on this trait than Asians (.634). The effect size for Psychoticism on the EPQ-R (d = .229), indicates that Pasifika tended to score higher on this trait than Asians. The effect size for the Lie scale on the EPQ-R (d = -.331) indicates that Asians tended to score higher on this trait than Pasifika (see Figure 6.3.4).

The effect size for Responsibility on the GPP-I (d = .235) indicates that Pasifika tend to score higher on this trait than Asians (see Figure 6.3.6). Finally, the effect size for Emotional Stability scale on the GPP-I (d = .219) indicates that Pasifika tended to score higher on this trait than Asians (see Figure 6.3.7).

Table 6.3.5
T-test and d statistic values of scale scores for Mâori and Asian.

	Levene's Test		t-test for Equality of Means			
	F	Sig.	t	df	Sig. 2-tailed	d Statistic
EPQ-R Scales						
P	0.002	0.962	0.542	82	0.589	0.139
			0.517	29.63	0.609	
E	0.578	0.449	0.936	82	0.352	0.240
			1.014	36.555	0.317	
N	0.371	0.544	-0.566	82	0.573	-0.145
			-0.51	27.471	0.614	
L	0.072	0.79	-1.562	82	0.122	-0.400
			-1.541	31.151	0.133	
GPP-I Scales						
A	2.062	0.155	0.772	82	0.442	0.197
			0.635	24.777	0.531	
R	0.067	0.796	-0.624	82	0.534	-0.160
			-0.609	30.585	0.547	
ES	0.097	0.756	0.086	82	0.932	0.022
			0.083	30.271	0.935	
S	0.214	0.645	0.445	82	0.658	0.114
			0.417	28.923	0.68	
C	0.184	0.669	-0.14	82	0.889	-0.036
			-0.126	27.505	0.901	
O	0.001	0.969	-0.606	82	0.546	-0.155
			-0.636	34.41	0.529	
PR	0.594	0.443	0.254	82	0.8	0.065
			0.277	36.887	0.784	
V	0.528	0.47	0.756	82	0.452	0.194
			0.806	35.465	0.426	

6.3.2 Overall Ethnic Differences on Personality Traits

Discriminant analysis was expected to yield some additional precision and depth beyond the output of *t*-tests and effect size scores. DSC was used to scrutinise how ethnic groups or gender groups differ from one another, in the total space of personality variables.

Multivariate analyses were conducted to further explore how ethnic groups differ on the personality scales in the EPQ-R and the GPP-I. Discriminant analyses were used to examine the contribution of personality variables to differences between ethnic groups. The analyses defined discriminant functions that were further checked for eigenvalue as well as for the structure matrix. The discriminant functions were expected to show how personality traits collectively 'discriminate', or fail to discriminate, pre-defined groups of cases in our data set.

Table 6.3.6

T-test and d statistic values of scale scores for Pasifika and Asians.

	Levene's Test		t-test for Equality of Means			
	F	Sig.	t	df	Sig. 2-tailed	d Statistic
EPQ-R Scales						
P	0.204	0.654	0.695	35	0.492	0.229
			0.706	34.966	0.485	
E	0.843	0.365	0.068	35	0.946	0.022
			0.067	31.6	0.947	
N	2.321	0.137	-0.134	35	0.894	-0.044
			-0.139	33.215	0.891	
L	0.659	0.422	-1.003	35	0.323	-0.331
			-1.011	34.85	0.319	
GPP-I Scales						
A	0.564	0.458	0.224	35	0.824	0.074
			0.229	34.326	0.82	
R	0.113	0.738	0.711	35	0.482	0.235
			0.712	34.258	0.481	
ES	0.026	0.872	0.663	35	0.512	0.219
			0.667	34.705	0.509	
S	0.498	0.485	-0.109	35	0.914	-0.036
			-0.108	32.969	0.914	
C	0.061	0.806	-0.476	35	0.637	-0.157
			-0.478	34.667	0.635	
O	1.924	0.174	0.287	35	0.776	0.095
			0.296	33.231	0.769	
PR	0.156	0.695	1.923	35	0.063	0.634
			1.905	32.428	0.066	
V	3.285	0.078	-0.338	35	0.737	-0.112
			-0.351	32.587	0.728	

DSC was performed with definitions of criterion groups New Zealand European (Pâkeha) vs. Mâori, NZ European vs. Pasifika, NZ European vs. NZ Asian, Mâori vs. Pasifika, Mâori vs. NZ Asian, and Pasifika vs. NZ Asian. The variables used to discriminate each pair of contrasted groups were defined as the EPQ-R trait scores and then (separately) the GPP-I trait scores.

Based on both the scrutiny of eigenvalues for the discriminant functions and the estimate of the ratio of cases "correctly classified" in discrimination, DSC did not result in any significant results for the EPQ-R traits. Results of these analyses are available on request (they are not reported in the thesis).

In analyses conducted with GPP-I traits, again the majority of discriminant functions showed minimal eigenvalues, and did not classify cases satisfactorily, pointing to the insufficient potential of these variables to establish a group contrast. These discriminant functions are not reported in the thesis (numerical findings are available on request).

There was one marginal exception, namely, the discriminant function for comparisons between Mâori and Pasifika over GPP-I personality traits. A selection of statistical findings from this analysis, including the structure matrix, eigenvalue of function, and checks on preconditions via F tests, are reported in Table 6.3.7 below.

While this result still shows a relatively low eigenvalue, it is higher than those produced in other group contrasts; further, this discriminant function correctly classified 67.2% of Mâori and 76.5% of Pasifika cases (the latter is the maximum correct classification produced in all analyses in the study).

Table 6.3.7.
Eigenvalue, canonical correlation, and Wilk's lamba results for Mâori-Pasifika group from Discriminant Analysis on the GPP-I.

Function	Eigenvalue	Canonical correlation	Wilksla mba	Chi square	df	*p*
1	.239(a)	.439	.807	16.089	8	.041

a First 1 canonical discriminant functions were used in the analysis.
b. Percentage

Table 6.3.8.
The Structure Matrix for Mâori-Pasifika group from Discriminant Analysis on the GPP-I

	Function 1
A	.432
R	.340
ES	-.248
S	.205
C	.171
O	-.128
PR	-.121
V	-.105

Table 6.3.9.
Standardized Canonical Discriminant Function Coefficients for Mâori-Pasifika group from Discriminant Analysis on the GPP-I

	Function 1
A	.498
R	1.074
ES	-.356
S	-.668
C	-.843
O	.336
PR	.716
V	-.736

6.4. Gender Differences on Personality Scales

In addition to the influence of ethnic group on personality traits gender-group comparisons were examined. The influence of gender was computed using the same analysis used for the examination of ethnic group influence. Thus, personality traits were examined as single variables performed by using a two-tailed *t*-test, Levene's test, and the *d* statistic. The variables examined for the comparisons were the EPQ-R and GPP-I trait scores; these variables all reach the interval level of measurement. Tables 6.4.1 presents group comparisons on each personality scale and display *t*-test, effect size, and the F ratio from Levene's test. The most important differences were considered to be differences that were significant and exceeding 0.2 standard deviations. In addition, Levene's F ratio reaching $p<.05$ were also deemed important.

Between males and females there were six significant gender group differences.
Across the GPP-I traits the difference of greatest magnitude was Emotional Stability, this difference was significant at $t(841) = 5.62$, $p<.01$, with males scoring higher than females ($d = .450$). Responsibility on the GPP-I had a significant difference between males and females $t(841) = -3.87$, $p<.01$, with males scoring higher than females ($d = -.310$). Vigor had a significant difference between males and females $t(841) = -3.71$, $p<.01$, with males scoring higher than females ($d = -.30$). Sociability on the GPP-I had a significant difference between males and females $t(841) = -2.60$, $p<.01$ with males scoring higher than females ($d = -.21$). Ascendency on the GPP-I had a significant difference between males and females $t(841) = 2.06$, $p<.05$, with females scoring higher than males ($d = .17$). Finally, Original Thinking on the GPP-I had a significant difference between males and females $t(841) = 1.89$, $p<.10$, with males scoring higher than females (.15).

Across the EPQ-R scales the difference of greatest magnitude between gender groups was Neuroticism, this difference was at $t(841) = -3.49$, $p<.01$, with females scoring higher than males ($d = -.285$). Psychoticism on the EPQ-R had a significant difference between males and females $t(841) = 2.54$, $p<.05$, with males scoring higher than females ($d = .21$).

Table 6.4.1
T-test and d statistic values of scale scores for Males and Females.

	Levene's Test		t-test for Equality of Means			
	F	Sig.	t	df	Sig. 2-tailed	d Statistic
EPQ-R Scales						
P	6.91	0.01	2.54	841.00	0.01**	0.21
			2.76	386.07	0.01	
E	0.00	0.97	-1.66	841.00	0.10	-0.13
			-1.69	342.76	0.09	
N	0.03	0.86	-3.49	841.00	0.00**	-0.28
			-3.44	327.35	0.00	
L	0.14	0.70	-0.96	841.00	0.34	-0.08
			-0.98	344.00	0.33	
GPP-I Scales						
A	1.12	0.29	2.06	841.00	0.04*	0.17
			2.14	355.90	0.03	
R	0.55	0.46	-3.87	841.00	0.00**	-0.31
			-3.88	335.60	0.00	
ES	0.16	0.69	5.62	841.00	0.00**	0.45
			5.58	330.92	0.00	
S	0.91	0.34	-2.60	841.00	0.01**	-0.21
			-2.51	315.92	0.01	
C	2.74	0.10	-1.22	840.00	0.22	-0.10
			-1.31	377.54	0.19	
O	2.60	0.11	1.89	841.00	0.06	0.15
			2.07	391.73	0.04	
PR	1.40	0.24	1.56	841.00	0.12	0.13
			1.50	314.44	0.13	
V	2.09	0.15	-3.71	841.00	0.00**	-0.30
			-3.55	312.31	0.00	

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

6.5. The Impact of Ethnic Status on Gender- and Age- relevant Patterns in Personality Assessment

Our aim (cf. Chapter 4) was to analyse data in order to locate any differences in personality traits indicating a disparity between two ethnic groups. After initially exploring the ethnic inter-group comparisons with personality traits the analysis then explored the influence of ethnic groups on the relationship between gender and age on personality traits. This examination was performed using multivariate analysis, with computation of alpha = .05. There were no overall significant results using Pillai's Traces, Wilks' Lambda, and Hotelling's Trace on the EPQ-R and GPP-I. On individual scales there were some significant interaction effects. Ethnic group and age had a significant interaction effect on Psychoticism on the EPQ-R $F(33) = 1.575$, $p<.05$, while ethnic group and gender had a significant interaction effect on Cautiousness on the GPP-I $F(3) = 2.377$, $p<.10$.

To further explore the relationship between ethnicity and gender on personality traits it was decided that again inter-group comparisons between males and females with a personality trait as a single variable would be performed by using a two-tailed *t*-test, Levene's test, and the *d* statistic. However, instead of computing this for the total number of cases this stage of the analysis was computed separately for NZ European and Mâori. The variables examined for the comparisons were the EPQ-R and GPP-I trait scores; as previously noted these variables all reach the interval level of measurement. Table 6.5.1 and 6.5.2 present the results of this analysis. Due to sample size level these were the only ethnic groups that were analysed.

For NZ Europeans there were three significant differences between males and females. Emotional Stability on the GPP-I had the largest difference with significant differences between males and females at $t(264) = 2.464$, $p<.05$, indicating that males tended to score higher on this trait than females ($d = .364$). Vigor had a significant difference between males and females at $t(264) = -2.188$, $p<.05$, indicating that females tended to score higher on this trait than males ($d = -.323$). Finally, Cautiousnes on the GPP-I had a significant difference between males and females at $t(264) = -2.160$, $p<.05$, indicating that females tended to score higher on this trait than males ($d = -.319$).

For Mâori there were two significant differences between males and females at $p<.10$. Ascendency on the GPP-I had the largest difference with significant differences between males and females at $t(54) = 1.937$, $p<.10$ indicating that males tended to score higher on this trait than females ($d = .535$). Sociability had a significant difference between males and females at $t(54) = 1.871$, $p<.10$, indicating that males tended to score higher on this trait than females ($d = .516$).

In addition to the two significant differences reported between males and females for Mâori the effect sizes on Table 6.5.2 indicates some additional differences. The effect size for Extraversion on the EPQ-R ($d = -.356$) indicates that females tended to score higher on this trait than males. The effect size for Personal Relations for Mâori on the GPP-I ($d = .290$) indicates that males tended to score higher on this trait than females. The effect size for Cautiousness ($d = -.262$) indicated that females tended to score

higher on this trait than males. Finally, the effect size for Responsibility for Mâori on the GPP-I (d = -.235) indicates that females tended to score higher on this trait than males.

6.6. Job Offer

The interrelationship between traits and Job offer was examined to determine how traits were related to Job Offer. The procedure was similar to that performed for the examination of bivariate relationships between traits, however, instead of two traits being compared at a time, traits were examined independently in relation to job offer. According to Steven's level of measurement trait scores were measured at the interval level of measurement, therefore, Pearson product-moment correlation coefficients (r) was chosen for this procedure. Table 6.6.2 shows the correlation Table of EPQ-R and GPP-I trait scores with job offer.

Table 6.5.1
T-test and d statistic values of scale scores between males and females for NZ European.

	Levene's Test		t-test for Equality of Means			
	F	Sig.	t	df	Sig. 2-tailed	D Statistic
EPQ-R Scales						
P	1.260	.263	.200	264	.841	0.030
			.219	108.257	.827	
E	.257	.613	-.478	264	.633	-0.070
			-.477	93.545	.634	
N	.090	.764	-1.125	264	.261	-0.166
			-1.155	97.247	.251	
L	1.160	.282	-.056	264	.955	-0.008
			-.060	103.490	.952	
GPP-I Scales						
A	3.577	.060	.157	264	.876	0.023
			.181	119.458	.857	
R	1.566	.212	-1.243	264	.215	-0.183
			-1.246	93.972	.216	
ES	1.562	.212	2.464	264	.014*	0.364
			2.320	86.401	.023	
S	1.160	.282	-2.188	264	.030*	-0.323
			-2.001	83.416	.049	
C	.453	.502	-.512	264	.609	-0.076
			-.540	101.738	.590	
O	1.073	.301	.963	264	.337	0.142
			1.021	102.454	.310	
PR	.023	.879	.704	264	.482	0.104
			.748	102.627	.456	
V	.399	.528	-2.160	264	.032*	-0.319
			-2.226	97.914	.028	

* Correlation is significant at the 0.05 level (2-tailed).

For the total number of cases in the study the largest correlation coefficient was -.231 and the smallest coefficient was .035, with a mean coefficient of ±.142 on the EPQ-R and ±.089 on the GPP-I. Table 6.6.2 displays six significant correlations between traits and job offers. However, the significant correlations found suggest only small degrees of shared variance among trait scores.

For NZ Europeans in the study the largest correlation coefficient was -.214 and the smallest coefficient was -.015, with a mean coefficient of ±.154 on the EPQ-R and ±.100 on the GPP-I. Table 6.6.2 again displays six significant correlations between traits and job offers. However, the significant correlations found suggested only small amounts of shared variance among trait scores.

Table 6.5.2
T-test and d statistic values of scale scores between males and females for Mâori.

	Levene's Test		t-test for Equality of Means			
	F	Sig.	t	df	Sig. 2-tailed	d Statistic
EPQ-R Scales						
P	3.391	.071	-.095	54	.925	-0.026
			-.106	53.733	.916	
E	1.878	.176	-1.289	54	.203	-0.356
			-1.438	53.813	.156	
N	.539	.466	-.746	54	.459	-0.206
			-.718	37.425	.477	
L	.088	.768	-.557	54	.580	-0.154
			-.565	44.217	.575	
GPP-I Scales						
A	.081	.777	1.937	54	.058	0.535
			1.988	45.670	.053	
R	.231	.633	-.850	54	.399	-0.235
			-.831	39.297	.411	
ES	.121	.729	-.095	54	.925	-0.026
			-.095	41.917	.925	
S	.273	.604	1.871	54	.067	0.516
			1.726	32.570	.094	
C	.344	.560	-.948	54	.347	-0.262
			-.977	46.198	.334	
O	.289	.593	.598	54	.553	0.165
			.624	47.874	.536	
PR	1.272	.264	1.053	54	.297	0.290
			.969	32.320	.340	
V	2.984	.090	.186	54	.853	0.051
			-2.226	97.914	.028	

* Correlation is significant at the 0.05 level (2-tailed).

Mâori in the study had a mean correlation coefficient of ±.116 on the EPQ-R and ±.223 on the GPP-I, with the largest coefficient of -.523 and the smallest coefficient

of .019. For Mâori there were some trait scores that suggest significant amounts of shared variance with job offer. The correlation coefficient between Vigor on the GPP-I and job offer suggests a negative relationship (-.523) at the significance level of $p<.01$. The correlation coefficient between the Ascendency on the GPP-I and job offer suggests a negative relationship (-.309) at the significant level of $p<.05$. Finally, the correlation coefficient between Neuroticism on the EPQ-R and job offer indicates a positive relationship (.320) at the significance level of $p<.05$.

For Pasifika in the study the largest correlation coefficient was .360 and the smallest coefficient was .048, with a mean coefficient of ±.283 on the EPQ-R and ±.163 on the GPP-I. For Pasifika there were no significant correlations between traits and job offer. However, this is likely to be due to the sample size of Pasifika (N = 17), as there were two correlation coefficients that suggested shared variance between specific personality traits and job offer. The correlation coefficient between Neuroticism on the EPQ-R and job offer indicates a positive relationship (.360), while the correlation coefficient between Extraversion on the EPQ-R and job offer indicates a negative relationship (-.331).

For Asians in the study the largest correlation coefficient was -.436 and the smallest coefficient was .012, with a mean coefficient of ±.215 on the EPQ-R and ±.196 on the GPP-I. There was one significant correlation between traits and Job Offer. The correlation coefficient between Sociability on the GPP-I and job offer indicates a negative relationship (-.436) at the significant level of $p<.05$.

In addition to analysing data in order to locate any differences in personality traits that indicate a disparity between two ethnic groups, the aim of the study was also to determine whether differences on personality traits among ethnic groups impact on selection outcomes. This stage of the analysis examined whether any differences in personality traits indicate a disparity between individuals selected for employment and individuals not selected for employment (i.e. job offer variable). To examine this aim inter-group comparisons with a personality trait as a single variable were performed by using a two-tailed *t*-test, Levene's test, and the *d* statistic. The variables examined for the comparisons were the EPQ-R and GPP-I trait scores; these variables all reach the interval level of measurement. Table 6.6.2 presents group comparisons

on each personality trait and displays *t*-test, effect size, and the F ratio from Levene's test. The most important differences were considered to be those that were significant and exceeding 0.2 standard deviations.

Across the EPQ-R three of the traits had significant differences. Psychoticism was significantly different between groups $t(943) = -3.69$, $p<.01$, indicating that those not offered employment scored higher on this scale ($d = -.25$). Extraversion was also significantly different between groups $t(943) = 5.39$, $p<.01$, indicating that those offered employment scored higher on this scale ($d = .36$). Finally, Neuroticism on the EPQ-R was significantly different between groups $t(943) = -6.23$, $p<.01$, indicating that those not offered employment scored higher on this scale ($d = -.41$).

Table 6.6.1
Correlations between personality traits and Job offers.

	Total cases	NZ European	Māori	Asian	Pasifika
EPQ-R Scales					
P	.127**	.208**	-.094	-.250	.261
E	-.173**	-.168**	-.019	-.331	-.191
N	.231**	.174**	.320*	.360	.284
L	.035	.065	.032	.191	.123
GPP-I Scales					
A	-.202**	-.182**	-.309*	-.137	-.211
R	-.042	-.071	-.080	.048	.111
ES	-.009	-.015	.066	-.216	.281
S	-.056	-.048	.009	.131	-.436
C	-.020	-.098	-.044	.236	-.211
O	-.116**	-.099	-.058	-.216	-.146
PR	-.043	-.069	.103	.299	.012
V	-.224**	-.214**	-.523**	-.117	-.156
N =	283	283	63	20	17

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Across the GPP-I three traits revealed significant differences. Ascendency was significantly different between groups $t(943) = 6.33$, $p<.01$, indicating that those offered employment scored higher on this scale ($d = .42$). Original Thinking was also significantly different between groups $t(943) = 3.57$, $p<.01$,. indicating that those offered employment scored higher on this scale ($d = .24$). Finally, Vigor on the GPP-I was significantly different between groups $t(943) = -6.23$, $p<.01$, indicating that those not offered employment scored higher on this scale ($d = -.41$).

Table 6.6.2

T-test and D statistic values of scale scores for Job Offer

	Levene's Test		t-test for Equality of Means			
	F	Sig.	t	df	Sig. 2-tailed	d Statistic
EPQ-R Scales						
P	0.20	0.65	-3.69	943	0.00**	-0.25
			-3.70	816.44	0.00	
E	21.94	0.00	5.39	943	0.00**	0.36
			5.20	705.84	0.00	
N	0.12	0.73	-6.23	943	0.00**	-0.41
			-6.18	789.50	0.00	
L	0.33	0.56	-1.07	943	0.29	-0.07
			-1.06	787.43	0.29	
GPP-I Scales						
A	21.29	0.00	6.33	943	0.00**	0.42
			6.03	672.47	0.00	
R	4.07	0.04	1.29	943	0.20	0.09
			1.25	718.02	0.21	
ES	16.48	0.00	0.28	943	0.78	0.02
			0.27	685.09	0.79	
S	12.95	0.00	1.71	943	0.09	0.11
			1.64	698.82	0.10	
C	15.51	0.00	0.63	942	0.53	0.04
			0.61	708.66	0.55	
O	10.64	0.00	3.57	943	0.00**	0.24
			3.51	764.48	0.00	
PR	12.03	0.00	1.32	943	0.19	0.09
			1.28	730.36	0.20	
V	10.80	0.00	7.05	943	0.00**	0.47
			6.82	712.57	0.00	

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

7. Discussion

This study has presented the findings of one of the first independent evaluations on the potential impact of demographic factors such as ethnicity on personality traits, such as those measured by the EPQ-R and the GPP-I. The current study had six specific research questions that it aimed to answer:

(i) Analyse New Zealand data to find potential relationships between personality and ethnicity. This included comparing ethnic groups along relevant personality variables (e.g. by means of two-tailed *t* tests utilising the means as well as effect size differences and standard deviations of ethnicity defined subsets among the participants), as well as an analysis of the impact of ethnic status on gender - and age - relevant patterns in personality assessment.

(ii) Provide further insight as to whether statistical comparisons between ethnic groups reflect "real" similarities and differences.

(iii) Provide information for New Zealand ethnic groups on other models of personality besides the FFM that are currently used in personality research and personnel selection.

(iv) Check whether any relationships between personality and ethnicity can be shown to be pervasive, irrespective of choice of a (systemic) model.

(v) Examine whether differences observed on personality assessment scores significantly impact on employee selection outcomes.

(vi) Provide further data on the relationship between personality and ethnicity for New Zealand sample groups.

This discussion has been divided into three separate chapters. Chapter 7 discusses the analysis of the correlation matrices analysed, then compares the trait scores among

ethnic groups and their consistency across personality inventories and models, then compares the variance of trait scores among ethnic groups and also the impact of ethnicity on gender and age relevant patterns on the EPQ-R and the GPP-I, and finally discusses the utility of the personality differences among ethnic groups in personnel selection settings. Chapter 8 outlines the limitations with suggestions for future research and then finally chapter 9 presents the conclusions of the research findings.

7.1. Generalisation issues

Empirical research in organisational psychology as well as psychology research conducted in the armed forces is subject to the same principles of corroboration and extrapolation as other empirical work in psychology. Ideally a precisely designed study would have participants that form a sample from a tightly defined population (Frankfort-Nachmias & Nachmias, 1996). The population would be defined prior to the sampling strategy and based on the goal of the research study. The sampling strategy and sampling procedure, again assuming a "perfectly" executed study, would follow the requirements set out in quantitative research methodology, especially in sampling theory (Hansen, Hurwitz, & Madow, 1993; Sampath, 2001; Singh, 2003). The nature of the sample will determine what can be inferred statistically from the findings, obtained through any techniques of inferential statistics.

Given the nature of the current study, with its divergence from the "ideal", it is important to address the restrictions upon generalisation (Lazerwitz, 1968). The restrictions also imply to specific "readings" (interpretations) of several statements made that follow in later sections of this discussion.

The current study is based on data accessed from a confidential selection process within an organization (cf. section 4 & 5). Due to the nature of the data the author could not determine in advance or change the nature of the assessment process. Further, this resulted in the author not having any impact on who was assessed, and whose data were recorded, retained, and/or released. These characteristics of the data used in the current study, resulted in the sampling process starting with a delimiting

population and thus, employing a standard procedure of sampling (e.g. probability sampling, Kalton, 1983; Lazerwitz, 1968; Stuart, 1987) was not viable in this study.

The thesis had to adopt and accept the participants and the set of assessment data kindly provided by the NZ Army. Because of the data characteristics it is important to realise that, although organisational psychology studies may opt to describe loosely recruited groups of participants as "samples", the participant groups in the current study are not samples under the strict methodological interpretation of the term. The main reason for this is that these groups are not random samples (Frankfort-Nachmias & Nachmias, 1996; Sampath, 2001). Therefore, their status vis-à-vis the population, the level of representativity, cannot be precisely stated given that a pre-existent definition of a population to sample from was not there.

Due to the sampling issues highlighted above it is important to note that we cannot generalise with any certainty the findings obtained from the present study to broader populations within New Zealand such as civilian employees/applicants or the entire adult New Zealand population. However, by examining some of the characteristics of the participants in the present study there are some observations that may allow at least some degree of external validity for the current research findings. While it is noted that the proportions for age and gender of the participants analysed in the present study are not consistent with the New Zealand populations there is some indication that the proportion of the four ethnic groups analysed in the present study appears to be proportional to New Zealand census data on ethnic groups in New Zealand (New Zealand census data 2006). This similarity allows some confidence that there are at least some similarities between the units of analysis in the present study and the population of interest (Lazerwitz, 1968).

It is worth reiterating that the present study is not alone in the present sampling issues. As stated previously, it is not uncommon in the realm of personality research, and in organisational studies, that authors report findings using samples of convenience (e.g. Barrick & Mount, 1992; Tett et al. 1992). Therefore, the present study in this respect is no different to some of the other research that has been conducted on personality and organisational psychology related issues.

With the issues raised with respect to the generalisation of the current study we need to consider to what extent the present research findings can be generalised. Firstly, it is highly likely that any research findings interpreted in the discussion apply to the total set of participants collected from archival New Zealand Army Officer Selection Board (OSB) data. Secondly, it is likely that the differences (or absence of), tendencies, or other research findings, apply to New Zealand army officers and cadets in general, at least in the last 5-10 years. Thirdly, we may have reasonable grounds to assume that the current research findings are true of personnel in New Zealand armed forces in general. Finally, some of the tendencies/findings may well be reproduced in civilian groups in New Zealand society, such as job applicants for civilian jobs or employed adult New Zealanders. If so, some of our findings may eventually be generalised to a wider population in New Zealand. This is well beyond the confines of the current study. Because of the sampling issues explained, a broad generalisation beyond armed forces personnel is without statistical grounds and will not be attempted.

7.2. Relationships among personality traits across ethnic groups

Before differences on personality traits were examined we first looked at the relationships among the personality traits within the EPQ-R and GPP-I. It also allowed some limited insight into applicability of theoretical assumptions in the personality models underlying the EPQ-R and GPP-I; namely, whether independence of traits in those models is shown by empirical findings on specific New Zealand ethnic groups that feature among participants in the present study. This also allowed for the observation as to whether the theories behind both the EPQ-R and GPP-I were applicable to the New Zealand ethnic groups analysed in the present study. This provided additional insight into the mean differences explored in later stages of the analysis, as we were able to determine whether mean score differences were associated with deviations in the interrelationships among personality traits.

7.2.1 EPQ-R

The overall pattern of correlations for the three traits in the EPQ-R were practically uniform across the ethnic groups among our participants, with minimal divergence among groups. These correlations were also relatively consistent with those reported in the EPQ manual (see Table 7.1). There was one exception, however, for Asians in the study as both correlations between Neuroticism and Extraversion and the Lie scale and Psychoticism were negatively correlated for these cases. Before we discuss the deviation that occurred for the Asian group, due to the primordial nature of the present study in investigating New Zealand ethnic groups on the EPQ-R, we will briefly reflect on the consistency of NZ Europeans, Mâori, and Pasifika correlations on the EPQ-R.

The results from the present analysis provide support for statistical independence of Psychoticism, Extraversion, Neuroticism, and the Lie scale across the New Zealand ethnic groups of NZ European, Mâori, and Pasifika (Because the data set received for the current analysis only has score-level information, rather than the item-level variables, it is not possible to replicate multivariate analyses showing factor structure in the item space of EPQ-R for the current participant groups. However, the correlations do suggest a level of relationship among the traits and scales that do invite the notion of statistical independence). More importantly the independence of the three traits in the EPQ-R suggests that what is being measured by the EPQ-R is consistent with those behaviours proposed by PEN (Eysenck three factor model, cf. Section 2.1.2.). The uniformity of matrices is a first, tentative indication that assessment systems can rely on tools developed on the basis of the PEN model at least for NZ European, Mâori, and Pasifika participants. This claim needs to be corroborated later on representative samples. Overall, it seems that assessors can be more confident that the behaviours proposed to be linked to Psychoticism, Extraversion, and Neuroticism are similar across NZ European, Mâori, and Pasifika participants.

In contrast, Asians in the present study had high correlations between Neuroticism and Extraversion. This indicates that for this ethnic group, personality traits are not independent but rather dependent on one another, with highly neurotic Asians likely to be highly introverted. The relationship suggested by the correlation between

Neuroticism and Extraversion for Asians is contrary to the theory behind PEN (cf. chapter 2), as the theory proposes that the three dimensions of personality are orthogonal. This suggests that for Asians in New Zealand, Neuroticism and Extraversion may not measure the same psychological constructs proposed by PEN.

The notion that, for Asians Extraversion and Neuroticism may not measure the same psychological constructs that are proposed by PEN, is also indicated by their relationship with traits measured by the GPP-I. For Asians, Extraversion and Neuroticism correlations with Vigor and Original Thinking are slightly stronger than correlations reported for other ethnic groups in the study and also correlations proposed in the GPP-I manual (see Table 7.2). Further, Neuroticism also shows slight deviations in correlations from other ethnic groups and those proposed in the GPP-I manual for Emotional Stability, Sociability and Ascendency, while Extraversion also shows slight deviations for Responsibility and Cautiousness. These correlations do suggest that for Asians in the present study Extraversion and Neuroticism may measure different psychological constructs than those proposed by PEN. More specifically the correlations identified above suggest that, for Asians, Extraversion and Neuroticism may be associated with slightly different behaviours than other ethnic groups.

For Asians this finding provides preliminary support for the notion that high extraverts and low neurotics may have a stronger association with intellectual curiosity, the tendency to think about new ideas, the tendency to work rapidly, and work productivity. While high introverts and high neurotics may have a stronger association with the dislike of working on difficult activities, non interest in thought provoking activities, and prefer settings of slower pace than other ethnic groups. In addition, the correlations also suggest that Extraversion for Asians is less associated with behaviours such as sticking to assigned jobs, preservation and determination, and the tendency to consider very carefully before making decisions compared to other ethnic groups.

Further, these correlations suggest that high neurotics may have stronger associations with restrictions in social settings, lack of gregariousness, the dependency on others for advice, and passive roles in group settings. While low neurotics may have a

stronger association with the enjoyment of working with people, gregariousness and Sociability, the tendency to make independent decisions, and active participation in group settings compared to other New Zealand ethnic groups. The change in the psychological constructs for Extraversion and Neuroticism suggested here, give rise to the possibility that the theoretical rationale behind the positive consequences of measuring these traits for Asians are not supported, as evidence suggests that the behaviours associated with these traits are not consistent for this group compared to other ethnicities.

Table 7.1
Correlations among scales for 500 Men and 500 Women from the EPQ-R manual.

	P	E	N	L	
P	-	.23	.19	-.34	Males
E	.14	-	.02	-.32	
N	.06	-.07	-	-.25	
L	-.16	-.19	-.26	-	
	Females				

Note. Eysenck & Eysenck (1991)
a. Sample information not supplied

For Asians, Psychoticism and the Lie scale also appeared to have a stronger relationship compared to other ethnic groups in the study. While this relationship is commonly observed on the EPQ-R (Table 4.1) it is slightly stronger than the commonly observed relationship, which is usually about <-.34. The relationship may suggest that for Asians low Lie scale scores are associated with behaviours such as insensitivity to other people's feelings, a disregard for danger, and the enjoyment of upsetting others. At first glance this is perhaps not what one would expect when considering what is supposed to be measured by the Lie scale and the Psychoticism trait, respectively. However, for Asians, Neuroticism and the Lie scale appear to be more independent (-.008) compared to the small correlation observed for other ethnic groups. This suggests that the Lie scale is measuring not the tendency to fake good but the tendency to behave in a socially conforming manner (cf. chapter 5.2). Thus, within this context it is possible to presume that low Lie scores would be associated with high psychotic individuals as it would suggest that individuals both show a

disregard for social conformity and are hostile, unconventional, and have a disregard for other peoples feelings.

Overall, Psychoticism is highlighted by correlational findings in the present analysis. Its correlations with GPP-I traits are less uniform across ethnic groups than the correlational pattern of Extraversion and Neuroticism with the same GPP-I traits. Psychoticism had low correlations with the majority of the traits on the GPP-I. The only correlations that indicated significant amounts of shared variance were with Cautiousness for all groups, the Responsibility trait for Mâori, the Original Thinking trait for Pasifika and Personal Relations for Asians. The correlations with the traits listed do appear to be logical associations that Psychoticism may have with the GPP-I. However, the variance of these correlations among ethnic groups suggests that it is possible that Psychoticism represents a slightly different construct for the groups listed. This finding is important as such variations for Psychoticism among ethnic groups suggests that the theoretical rationale behind the positive consequences of measuring this trait are not supported. In other words, this finding suggests that the theoretically related behaviours to Psychoticism may deviate among New Zealand ethnic groups.

Discussed previously in section 2.1.2 and 5.2.2 the Psychoticism trait is much more controversial than the other two personality traits proposed by PEN (Block, 1977; Bishop, 1977; Eysenck, 1985; Eysenck & Eysenck, 1991; Kline, 1993) and as a result the exact nature of Psychoticism is less clear. From the present findings it is possible that some of the inaccuracy and inconsistency when exploring the Psychoticism trait can be explained by the ethnicity of individuals. However, from the data presented in the current study such interpretations can only be speculative.

We should note that the findings for Asians are not consistent with the vast amount of research conducted on the EPQ and EPQ-R assessments, as the majority of this research has found consistent independent factor structures across many different ethnic groups (Eysenck & Eysenck, 1991; Barrett & Eysenck, 1984). Furthermore, research conducted using similar samples (Chinese) has also not found this relationship between Extraversion and Neuroticism (Eysenck & Chan, 1982). Therefore, we may conclude that the ability to generalise from this finding is

restricted. Yet it can still be argued that personality research conducted on Asian populations is one of the more inconsistent areas of study with culture specific factors being observed in many studies examining the factor structure of personality for different personality models (e.g. Cheung, Leung, Zhang, Sun, Gan, Song, & Xie, 2001). Overall, what we can gather from the correlations among the traits measured by the EPQ-R is that the Asian group in the present study show slight differences to the other ethnic groups and past research, hence, suggesting that the Asian group show inconsistent relationships among EPQ-R personality traits to what has previously been observed.

Table 7.2
Correlation matrix between the EPI scales and the GPP-I scales from the GPP-I manual.

	A	R	ES	S	C	O	PR	V
Civil Service Employees (N = 123)								
Extroversion	.58	-.06	.10	.57	-.40	.24	-.13	.40
Neuroticism	-.40	-.22	-.53	-.28	-.05	-.31	-.23	-.34

Note. Gordon (1978)
a. Participants were 123 Civil Service employees.

7.2.2 GPP-I

The relationships among personality traits on the GPP-I show slightly more variation across ethnic groups than relationships among EPQ-R traits did. However, overall there was considerable similarity among ethnic groups and relationships proposed by the GPP-I manual using a management sample (see Table 4.3). Pasifika respondents were the only ethnic group in the current study whose scores produced a correlational pattern quite different to other ethnic groups, with Ascendency, Responsibility, and Emotional Stability being highly correlated with one another. These high correlations among the three traits suggest a much stronger relationship than reported for other ethnic groups in the study. In addition, these correlations are much higher than the relationship proposed among the three traits in the GPP-I manual (see Table 4.3). Although Gordon (1978) proposes that correlations among the GPP-I traits are likely to be found the high magnitude of correlations among Ascendency, Responsibility and Emotional Stability do question the validity of these traits for Pasifika. The

correlations reported suggest very high levels of shared variance among the three traits suggesting that they are measuring functionally the same or a very similar psychological construct. Furthermore, this finding indicates that the three traits for Pasifika do not provide any additional incremental validity from one another.

Consistent with the correlations reported within the GPP-I for Pasifika, correlations between the EPQ-R and GPP-I for Pasifika also report deviations for Ascendency, Responsibility, and Emotional Stability compared to other ethnic groups. These correlations deviate with other ethnic groups in the study and those proposed in the GPP-I manual (Table 7.2). For Pasifika in the present study both Ascendency and Responsibility correlated strongly with the Lie scale on the EPQ-R, which was not consistent with other ethnic groups. Emotional Stability also appeared to show a relationship with the Lie scale, however, this was not significant due to sample size. In addition, although not significant, Responsibility also appeared to show a relationship with Psychoticism while Emotional Stability also correlated with Extraversion.

The deviations for Pasifika on the traits of Ascendency, Responsibility, and Emotional Stability question the validity of the GPP-I for Pasifika samples. There were stronger relationships among Ascendency, Responsibility, and Emotional Stability than in other applicant groups. This may put the theoretical rationale for the use of these scores into question. The findings also question the link between these traits and overt behaviours that could be predicted. This finding may gain some importance given the lack of empirical research examining the factor structure of the GPP-I across cultures. The lacuna is noticeable if GPP-I studies are compared to cross-cultural research exploring factor structure in tools based on other models of personality (e.g. FFM, PEN, 16PF). The lack of research examining different culture and ethnic groups means that the present study findings that Ascendency, Responsibility, and Emotional Stability are highly correlated for Pasifika, suggests that further research is needed. This finding presents the need for caution for Pasifika ethnic groups when interpreting the GPP-I (notably Ascendency, Responsibility, Emotional Stability) until further research on this ethnic group is conducted.

Table 7.3
Correlations among the GPP-I scales from the GPP-I manual.

	A	R	ES	S	C	O	PR	V
Manager Sample (N = 218)								
A	-	-.24	-.11	.68	-.30	.39	.17	.29
R		-	.51	-.19	.44	.12	.14	-.01
ES			-	-.37	.34	.11	.31	-.13
S				-	-.30	.17	.18	.19
C					-	-.26	.23	-.27
O						-	.01	.19
PR							-	-.23
V								-

Note. Gordon (1978) p. 22.
a. Management sample obtained from a USA Public Utility centre.

In summary the findings from this stage of the analysis suggest generally good consistency among NZ Europeans, Mâori, Pasifika, and Asians for the interrelationships among personality traits measured for both the EPQ-R and the GPP-I. In addition, the pattern of relationships among trait variables revealed in the current study also appears to be consistent with correlational patterns reported in assessment manuals for the EPQ-R and the GPP-I. However, there are some deviations from this pattern for Asian participants with higher than usual correlations found between Extraversion and Neuroticism and between the Lie scale and Psychoticism on the EPQ-R. Further, deviations to this pattern are also present for Pasifika participants with higher than usual correlations among Ascendency, Emotional Stability, and Responsibility on the GPP-I. This suggests that the trait scores for Asians on the EPQ-R and Pasifika on the GPP-I are perhaps less generalisable than the other ethnic groups in the present study. In addition, these deviations for the Asian and Pasifika group question the theoretical rationale for the positive consequences for measuring the specific traits that deviate, as such deviations question their theoretic link with the descriptive behaviours they are proposing to measure.

7.3. Inter-group Differences regarding Personality Traits

We now turn to the interpretation of the results for the major aim of the present study to examine the presence or absence of differences in personality trait scores among New Zealand ethnic groups. Two different approaches were used to analyse this research aim. First, a variable by variable approach viewing each personality construct independently. The personality variables were analysed collectively focusing on personality traits simultaniously.

While the discussion will include previous research, this is limited to predominantly Packman et al. (2005) due to the lack of appropriate sample comparisons. Each trait difference observed will also be analysed for consistency with the conceptually related traits between the EPQ-R and GPP-I. The traits chosen as conceptually similar across different inventories were based on pre-existing knowledge of the behaviours measured by these traits and also correlations observed between the EPQ-R and the GPP-I in the present study (Table 6.2.3.). The interpretation of the consistency between the EPQ-R and the GPP-I will then be discussed in the preceeding section.

7.3.1. EPQ-R

Overall, there was little variance among personality trait scores across ethnic groups on the EPQ-R. The only consistent difference among ethnic groups was on the Lie scale with Asians tending to score higher than other ethnic groups in the study.

Pasifika was the only ethnicity to show a difference, which occurred when compared with NZ Europeans (-.267) and Asians (.229). Both these differences indicated that Pasifika scored higher on this trait than NZ European and Asian. This suggests that Pasifika as a group have higher levels of Psychoticism in their personality than NZ Europeans and Asians. However, according to Cohen's standard these values suggest only a small difference between groups.

The difference observed for Pasifika when compared to NZ Europeans and Asians is not consistent with any of the major traits on the GPP-I that are conceptually similar and correlated with Psychoticism (i.e. Cautiousness and Responsibility). This finding also appears to be inconsistent with past research using the EPI inventory. Church and Lonner (1998) found that relative to a United Kingdom sample Chinese generally scored higher for Psychoticism. Although the populations in the present study are slightly different, we would have expected similar results if we consider McCrae (2002) findings, which suggested that countries in similar geographical locations generally have similar mean trait scores. Therefore, the outcome in the present study that Pasifika score higher than Asians and Asians do not score higher than NZ Europeans is not consistent with Church and Lonner (1998).

Extraversion

Mâori were the only group that showed a tendency to score higher than other groups for Extraversion, this difference rose when they were compared with Asians (.240). This suggests that Mâori in the present study tended to have higher levels of Extraversion than Asians. According to Cohen's standard these values only suggested a small difference.

The finding that Mâori scored higher than Asians was only mildly consistent with the conceptually similar Original Thinking trait on the GPP-I, as Mâori on this trait showed a small tendency to score lower than other ethnic groups (i.e. NZ Europeans). This finding is also contrary to past research investigating Mâori and Asians. Research using the 15FQ+ based on the FFM comparing Mâori to other ethnic groups suggests that Mâori have a small tendency (although effect sizes are close to suggesting no difference) to score lower than NZ Europeans and Pasifika (Packman et al. 2005). This difference also appears to be inconsistent with past research using the EPQ as Chinese populations in Asia have been found to score higher than other ethnic groups for Extraversion (Church & Lonner, 1998). Given McCrae's (2002) finding that trait means are subjected to geographical areas, with countries close geographically having similar mean trait scores, we would anticipate that Asians in the present study would score higher for Extraversion. However, this is not the case, as Asians appear to be similar to other ethnic groups apart from Mâori who, as stated, score lower than Asians.

Neuroticism

The only difference that was observed in the present study was between Asians and NZ Europeans for Neuroticism (-.201), which suggested that Asians scored higher than NZ Europeans for this trait. According to Cohen's standard this value only suggested a small difference.

This difference was not consistent with the conceptually similar trait of Emotional Stability and only consistent with Ascendency on the GPP-I, as this trait showed higher NZ Europeans scores compared with Asians. This difference however, did show some level of consistency with past research comparing NZ European groups with other ethnic groups. Packman et al. (2005) found that NZ Europeans tended to score lower compared with both Pasifika and Mâori on the similar trait measured by the 15FQ+ the Neuroticism trait based on the FFM and higher on the Emotionally Stable trait also measured on the 15FQ+.

Lie scale

The EPQ-R mean scores for the Lie scale shared a large amount of variation in the present study, as five of the six mean group comparisons had effect sizes above .20. The difference of greatest magnitude on this scale was between NZ Europeans and Asians, with an effect size of -.58 suggesting Asians scored higher than NZ Europeans on this trait. Overall, Asians on the Lie scale scored higher than all other ethnic groups in the study with effect sizes ranging from .331 (Asian-Pacific Islander group) to .58 (NZ European-Asian group).

The effect size observed between Asians and NZ Europeans is of great interest, as in cognitive ability research effect size scores at around .50 often indicate differences with the ability to cause adverse impact for a particular group (Sackett, Schmitt, Ellingson & Kabin, 2001). Furthermore, the Lie scale on the EPQ-R is used for measuring the tendency for respondents to fake good and therefore when high scores are present, interpreters of the test often question the overall results of the personality inventory. In worst case scenarios this may lead to respondent scores being disregarded and not considered for the purpose they have been administered e.g. selection. The magnitude of these differences is accentuated because the difference

for Asians is with NZ Europeans who represent the majority of the present data set, and suggests that the deviation is likely to show up in overall data analysis.

The tendency for Asians to score higher on this scale than other ethnic groups and specifically NZ Europeans is similar to past research comparing the similar ethnic samples of Chinese and Europeans in the United Kingdom. In these studies Chinese, like Asians in the present study, have scored higher than Europeans in the United Kingdom (Eysenck & Chan, 1982). Therefore, this finding suggests that Asians, like other regions of the world, score higher as a group on the Lie scale measured by the EPQ-R.

For the Asian sample the correlations between the Lie scale and Neuroticism suggest that they are independent from one another (see Table 6.1.1). This suggests that the Lie scale for Asians is more likely to reflect that there is more effort made to follow examples of well behaved, socially conforming individuals, rather than the tendency to fake good (cf. Section 5.2.1). Therefore, the potential for the higher scores for Asians on the Lie scale to discredit their responses may appear to be unjustified when measuring personality traits in Asian respondents.

7.3.2 GPP-I

The GPP-I showed more variance in mean personality trait scores among ethnic groups relative to the EPQ-R. The traits that showed the largest amounts of mean group variance were Personal Relations and Vigor. There was only one trait that did not report mean variance among ethnic groups this was for Cautiousness.

Ascendency

There was little difference among ethnic groups for Ascendency. The only difference observed was between NZ Europeans and Asians (.267), suggesting that NZ Europeans scored higher than Asians. This suggests that NZ Europeans in the present study tended to have higher level of Ascendency than Asians. However, according to Cohen's standard this value only suggested a small difference.

The difference between NZ Europeans and Asians is one that is not present in the conceptually and correlated trait Extraversion, measured by the EPQ-R, however, this

difference is present for the related trait of Neuroticism measured by the EPQ-R. Thus, this difference between NZ Europeans and Asians shows some consistency between the two inventories in the present study. However, this difference does appear to be contrary to past research analysing NZ Europeans. Packman et al. (2005) produced findings that suggest that NZ Europeans have close to the same levels of the conceptually similar Self-sufficient trait measured by the 15FQ+ and possibly have lower scores than other ethnic groups (i.e. compared to Mâori). Furher, the tendency for NZ Europeans to score lower than other ethnic groups on this trait is not consistent with Packman et al. (2005) on any trait measuring social aspects of personality.

Responsibility

NZ European, Mâori, and Asian ethnic groups showed little variance in mean scores among each other for Responsibility. However, Pasifika showed a tendency to score higher than all other ethnic groups in the study. Effect sizes suggested that Pasifika scored higher than Mâori (-.403), NZ Europeans (-.24), and Asians (-.235). According to Cohen's standards these effects sizes were small apart from the comparison between Pasifika and Mâori, which was closer to the moderate range.

The tendency for Pasifika to score higher than other ethnic groups for Responsibility is contradictory to findings for Psychoticism on the EPQ-R, which is conceptually similar to Responsibility and also correlated for the total number of cases in the present study. However, the tendency for Pasifika to score higher on this trait is consistent with Packman et al. (2005) who found that a similar tendency for Pasifika to score higher than Mâori and NZ European on a similar trait measured by the 15FQ+, Self-Discipline and Neuroticism based on the FFM.

Emotional Stability

NZ Europeans were the only group to show a difference for Emotional Stability, with an effect size of .21 when compared to Asians and .25 when compared with Pasifika. Both these differences indicated that NZ Europeans scored higher on Emotional Stability than Asians and Pasifika. According to Cohen's standard these effect sizes only suggest a small difference between groups.

The small difference between NZ Europeans was not consistent with the conceptually related trait of Neuroticism on the EPQ-R. However, the tendency for NZ Europeans to score higher than Pasifika is consistent with previous research comparing the two ethnic groups. Packman et al. (2005) found that NZ Europeans tended to score higher than Pasifika on the similar trait measured by the 15FQ+, Emotionally-Stable and lower on the Neuroticism trait based on the FFM.

Sociability

NZ Europeans were the only group to show differences on this trait, with effect sizes of .248 when compared with Asians and .206 when compared with Pasifika. Both these differences indicated that NZ Europeans scored higher on Sociability than Asians and Pasifika. According to Cohen's standard these effect sizes suggest a small difference among groups.

The difference between NZ Europeans and both Pasifika and Asians is not present on the conceptually related trait of Extraversion on the EPQ-R. This difference is also not observed in past research in analysing the differences between NZ Europeans and Pasifika on the theoretically related traits of Extraversion (based on the FFM) and Socially-bold measured by the 15FQ+ administered by Packman et al. (2005). However, Packman et al. did find some indication that NZ Europeans tended to score higher than other ethnic groups with a small difference found between NZ Europeans and Mâori.

Cautiousness

Among the New Zealand ethnic groups analysed in the present study there were no differences observed for Cautiousness. The largest difference was between Pasifika and Asians with an effect size of .154. Such low effect sizes among the ethnic groups suggested that this trait was very close to having invariance among the four ethnic groups.

The the consistency of the invariance among ethnic groups for Cautiousness appears mixed when comparing this finding to the differences observed for Psychoticism on the EPQ-R, as while this trait found a large amoung of invariance among ethnic groups it also found a small but observable difference between Pasifika and NZ

Europeans (-.267) and also with Asians (.229). Yet, the finding that Cautiousness had no variance across ethnic groups is consistent with previous research conducted using New Zealand samples. Packman et al. study found a similar occurrence with similar traits called Apprehensiveness and Conscientiousness measured by the 15FQ+. They found very small effect sizes among Mâori, NZ Europeans, and Pasifika, which suggested that the groups were close to having no variance among their mean scores.

Original Thinking

Mâori were the only group to show differences on this trait, with effects sizes of .29 when compared with NZ Europeans and -.244 when compared with Pasifika. Both these differences suggested that Mâori scored lower than NZ Europeans and Pasifika on this trait. According to Cohen's standard these effect sizes only suggested a small difference.

Although the difference between Mâori and NZ Europeans and Mâori and Pasifika are small the tendency for Mâori to score lower than other ethnic groups is consistent with differences among ethnic groups for Extraversion but not for Neuroticism on the EPQ-R. The Extraversion scale showed a small tendency for Mâori to score higher than Asians. Although the two traits have different comparisons, the two traits do show consistency with the result that Mâori are more extraverted and lower in Original Thinking, which is consistent with the conceptual understanding of the two traits.

The tendency for Mâori to score lower than NZ Europeans for Original Thinking also appears to be consistent with past research, when comparing the two groups. Packman et al. (2005) found that Mâori scored lower than NZ European on a similar trait measured by the 15FQ+, Intellectance. Although the lack of a significant *t*-test and the small effect size does not suggest a large difference, the difference observed is consistent with a previous pattern found when comparing the two groups.

Personal Relations

For the Personal Relations trait Pasifika showed relatively large differences compared to other ethnic groups in the study. Pasifika scored higher than all other ethnic groups in the study with an effect size of -.513 when compared with Mâori, -.583 when

compared with NZ European, and -.634 when compared with Asians (see Table 3.2.2, 3.2.4 and 3.2.6). All these scores suggested that Pasifika scored higher on this trait than other Asians, Mâori, and NZ Europeans. According to Cohen's standard these effect sizes suggest moderate differences between groups.

The finding that Pasifika scored higher on Personal Relations appears inconsistent with the conceptually related trait of Extraversion. In addition, this result also appears to be contrary to past research comparing this ethnic group to other New Zealand ethnicities. Packman et al. (2005) found that Pasifika scored higher than other ethnic groups (Mâori and NZ Europeans) on a similar trait measured by the 15FQ+, Suspicious. High scores on this trait are similar in nature to low scores for the Personal Relations trait. Thus, Packman's et al finding that Pasifika scored high on this trait is in direct disagreement with the present study's results. Packman et al. also found a difference between Mâori and NZ European, with Mâori scoring considerably higher than NZ Europeans, which is also not supported by the present study. Further, Packman et al. (2005) also found very little difference between Pasifika and other New Zealand ethnic groups on the similar trait of Extraversion measured by the 15FQ+.

Vigor

The Vigor trait was the other trait measured on the GPP-I that had large amounts of variance among mean scores. The largest difference was between NZ Europeans and Pasifika, with an effect size of .50 suggesting that NZ Europeans scored higher than Pasifika on this trait. Results also suggested that NZ European scored higher than Asians, with an effect size of .39. These results indicate that NZ Europeans have a tendency to score higher on the Vigor trait than other New Zealand ethnic groups. According to Cohen's standard the effect sizes between NZ European and Pacifika and NZ European and Asian were in the moderate range while the effect size between Mâori and Pacifika suggested a small difference.

The differences found among ethnic groups for Vigor are not consistent with findings for the conceptually related traits of Extraversion and Neuroticism on the EPQ-R. In addition, the difference between Mâori and Pasifika is also not consistent with past research. Packman et al. (2005) also found no difference on the conceptually similar

trait of Enthusiastic measured by the 15FQ+. However, it should be noted that Packman et al. did find some tendencies for NZ Europeans to score higher than other ethnic groups on this trait, with differences found when the ethnic group was compared to Mâori. In addtion, similar differences were also found with Mâori on the conceptually similar trait of Extraversion based on the FFM.

7.3.3. Overall Ethnic Differences on the EPQ-R and GPP-I

After the variable by variable approach personality variables were analysed collectively and focused on the overall pattern of personality traits simultaneously. Discriminant analysis was expected to yield some additional precision and depth beyond the output of *t*-tests and effect size scores. DSC was used to scrutinise how ethnic groups or gender groups differ from one another, in the total space of personality variables.

The DSC did not result in any significant results for the EPQ-R traits. Again the majority of discriminant functions showed minimal eigenvalues for the GPP-I traits, and did not classify cases satisfactorily. However, for the GPP-I traits there was one marginal exception, for comparisons between Mâori and Pasifika. This discriminant function correctly classified 67.2% of Mâori and 76.5% of Pasifika cases. However, the results also indicated that groups differed in their covariance (cf. Section 6.3.2) suggesting that analysis was in violation of multivariate assumptions.

Overall, the majority of non-significant findings from this stage of the analysis (variable by variable analysis and the collective approach) appears to be consistent with the personality differences observed from the variable by variable analysis. The variable by variable analysis shows that the impact of ethnicity on personality traits appears to be of a complex nature and the differences are not consistent across ethnic groups and personality traits. Rather exactly which trait(s) that were shown to differ between any two ethnic groups was largely dependent on the specific ethnic groups being compared and the personality traits being measured. Furthermore, the majority of differences found on single personality traits among ethnic groups were small by Cohen's standards (1988). Therefore, it would be expected that discriminant analysis would not produce significant results.

The only trait differences that were of moderate to large magnitude were the Personal Relations and Vigor traits measured on the GPP-I and the Lie scale measured on the EPQ-R. Therefore, this pattern of ethnic mean differences found in the present study is consistent with both past research finding differences of small magnitude (Goldberg et al. 1998; Guenole & Chernyshenko, 2005; Hough, 1998a; Ones &Anderson, 2002; Ones & Viswesvaran, 1998) and those studies finding significant differences among ethnic groups on personality inventories (Eysenck & Chan, 1982; Heuchert et al. 2000; Hoffman et al. 1985; Knapstein, 1970; Lujan & Zapata, 1983; Packman et al. 2005; Whitworth & Perry, 1990).

7.3.4. Stability of Ethnic Difference on Personality Traits

Both methods used to analyse the mean ethnic differences observed in the present study indicated that the differences on each personality trait were generally inconsistent with the conceptually related traits measured by the alternative inventory. From the variable by variable analysis on all the 12 traits and scales analysed in the present study. Only Ascendency-Extraversion and Original Thinking-Extraversion differences among ethnic groups had some degrees of consistency across the EPQ-R and the GPP-I, while Cautiousness showed a consistent pattern of invariance among ethnic groups across the two inventories. In addition, this lack of consistency between the EPQ-R and the GPP-I appears to also be constant when comparing the observed differences among ethnic groups to Packman et al.'s (2005) research findings, as only the Responsibility, Emotional Stability, Original Thinking, and Cautiousness trait differences appear to show some consistency with conceptually related traits measured by Packman et al.

The general lack of consistency between the two inventories and with past research, again using an alternative personality inventory, indicate some cause for concern for research examining the differences among ethnic groups on personality traits. The inconsistency suggests that any differences observed on personality traits among ethnic groups are likely to be specific to that personality inventory providing the measure of the trait. Further, because the nature of the two inventories used in the present study are based on different models of personality this indicates that the differences observed among ethnic groups are model dependent. In further support of this notion is the inconsistency with both personality inventories compared to the

other study investigating ethnic differences in New Zealand (i.e. Packman et al. 2005) using the 15FQ+, again based on another different personality model i.e. FFM.

The design of the present study does not allow us to determine the reason behind the lack of consistency between the two inventories based on the different models of personality. However, we may speculate as to the underlying reason for this difference. One explanation for the inconsistency is the phenomenon of measure-specific variance. This notion suggests that the reason for the lack of consistency among inventories on demographic differences is because personality scales that aim to measure a given trait, even when correlations are corrected for unreliability, average well below 1.00 (i.e. 100%). Consequently, this results in different measures of the similar personality traits not assessing the exact same definition of a given trait.

To further explain this issue One & Anderson (2005) suggest that personality measures can be conceptualised as arising from common factor variance and unique factor variance. This conceptualisation proposes that common factor variance differences are representative of the shared variance between personality measures, while unique factor variance is representative of the unshared variance between personality measures.

If we put this in the context of the present findings this suggests that the personality differences among the New Zealand ethnic groups are predominantly due to the presence of unique factor variance. We would expect some degree of unique variance present when comparing group differences across different personality inventories and personality models. Although it can be argued that, if these behavioral differences among ethnic groups are real, we would expect some level of similarity when comparing personality difference on conceptually similar traits across any two personality inventories based on different models.

This notion is a cause for concern in the present study, as this suggests that the results are predominantly due to the unique factor variance of the inventories, meaning that the differences that are observed are not real and therefore not valid. In other words, if this interpretation of the results is correct then according to the theory of measurement this suggests that the differences observed among the ethnic groups in the present

study are not a result of real differences but rather a result of artefacts created by the personality measures. Such possibilities are of major concern for the realm of personality assessment as it is important for the practice of any assessment procedure to distinguish between adverse consequences of an assessment that stem from valid descriptions to adverse consequences of an assessment that stem from invalid descriptions of individuals and groups (Messick, 1996). These observations are a concern for the personality traits of Personal Relations and Vigor measured by the GPP-I in the present study, as it is possible that such differences are not real differences among ethnic groups and thus not reflective of real behavioural tendencies.

From the present research findings, whether the underlying reason behind the lack of consistency between the two inventories in the present study is true or not is not clear. However, what does appear to be indicative of the present findings is that differences between and among ethnic groups are specific to the personality inventory and personality model they are observed on. Noted in section 2.1.2 such findings are a concern, as this suggests that the majority of the research findings analysing this issue that have used the FFM of personality (e.g. Goldberg et al. 1998; Heuchert et a. 2000; McCrae, 2001, 2002; McCrae & Terraciano, 2005; Ones & Anderson, 1998; Ones & Viswersan, 1998; Packman et al. 2005; Schmidt et al. 2007; Schmitt et al. 2008) may not be conceptually generalised to different personality models and inventories based from these research findings.

7.4. The Two Alternative Ways that Ethnicity could be Associated with Personality

In addition to investigating the direct impact of ethnicity on personality traits, the present study also aimed to analyse the two alternative ways ethnicity may influence personality. Firstly, the variance of personality traits among New Zealand ethnic groups was explored. Secondly, the impact of ethnicity on the relationship of age and gender on personality traits was analysed.

7.4.1 Variance of Personality Traits among Ethnic Groups

When comparing the standard deviations among ethnic groups (i.e. Levene's test) the present analysis did not find any significant difference among ethnic groups. This is a similar result to Ones and Anderson (2002) when they compared the variance of personality traits among gender groups. This finding is contrary to cognitive ability research on the variance of demographic groups (Feingold, 1994). While there were no differences observed in the current study because of the limited amount of research conducted in this area it is suggested that future research should continue to examine this issue, because as mentioned in section 5 this is an important aspect of ethnic differences in empirical assessment.

7.4.2 The Impact of Ethnic Status on Gender - and Age - relevant Patterns in Personality Assessment

To explore the effects between ethnicity and gender and age with respect to personality the present study used multivariate of analysis (MANOVA). This analysis resulted in no overall significant results using Pillai's Traces, Wilks' Lambda, and Hotelling's Trace on the EPQ-R and GPP-I. However, further exploration on the impact of ethnicity on gender indicated a small but observable impact of ethnicity on the relationship between gender and personality on both the EPQ-R and the GPP-I.

EPQ-R

Among the three traits and the Lie scale measured by the EPQ-R there was only one difference between NZ European gender groups and Mâori gender groups. There was a gender difference on Extraversion for Mâori while NZ Europeans had no difference among the EPQ-R measures. The difference between males and females for Mâori showed a tendency for Mâori females to score higher than Mâori males with an effect size of -.356. According to Cohen's standards this is a small to moderate difference.

The difference between males and females for Mâori on the Extraversion trait is perhaps a surprising one, because when examining the total number of cases there was only a very small difference observed on this trait between males and females. This deviation from the total number of cases for Mâori suggests that this finding is specific for Mâori and does not exist for other ethnic groups in the study.

GPP-I

Among the traits measured by the GPP-I there were four traits that showed observable differences between NZ European gender groups and Mâori gender groups. There was a gender difference on Ascendency for Mâori while NZ Europeans had no difference between gender groups. The difference suggested that Mâori males tended to score higher on this trait than Mâori females with an effect size of .535. According to Cohen's standard this effect size suggested a moderate difference between gender groups.

The tendency for Mâori males to score higher than Mâori females on this trait is consistent with the differences observed for Extraversion on the EPQ-R. This suggests that the small but observable difference between genders for Mâori is consistent across personality inventories and models. Further, the fact that NZ European gender groups did not differ on either trait suggests that there is a real trait difference between gender groups for Mâori while there appears to be real invariance on this trait for NZ Europeans.

For Emotional Stability there was an observable difference between genders for NZ Europeans while gender mean scores appeared to be invariant for Mâori. The difference between genders for NZ Europeans suggested that NZ European males tended to score higher than NZ European females with an effect size of .364. According to Cohen's standard this effect size suggests a small to moderate difference between groups. This difference between gender groups for NZ Europeans was not consistent with the conceptually related trait of Neuroticism on the EPQ-R. Thus, for NZ Europeans this difference appeared to only be observed for the GPP-I.

There was a gender difference on Sociability for Mâori and also for NZ Europeans. However, the direction of the relationship between gender and personality changed between Mâori and NZ Europeans, as Mâori males tended to score higher than Mâori females while NZ European males tended to score lower than NZ European females. The diametric results for gender differences between NZ European and Mâori suggest that there are substantially different relationships between gender and the Sociability trait measured on the GPP-I for NZ Europeans and Mâori.

Cautiousness on the GPP-I was another trait that had an observable difference between NZ Europeans and Mâori. For Mâori gender there was a small tendency for females to score higher than males on Cautiousness with an effect size of -.262, while for NZ Europeans gender groups were very similar. According to Cohen's standard this effect size suggests a small difference between groups. This small but observable difference between gender groups for Mâori is consistent with the conceptually similar traits of Extraversion on the EPQ-R. Thus, this suggests consistency across measures between gender groups for Mâori.

For Vigor there was an observable difference between genders for NZ Europeans while gender mean scores appeared to be invariant between genders for Mâori. The gender difference for NZ Europeans showed that females tended to score higher than males for Vigor with an effect size of -.319. According to Cohen's standard this effect size suggests a small difference between groups. The difference observed between gender groups for NZ Europeans was not consistent with the conceptually related trait Extraversion in the EPQ-R. This suggests that the difference observed between NZ Europeans is specific to the GPP-I inventory.

Overall, when comparing Mâori and NZ European gender groups the results indicate that Mâori have greater variance between gender groups than NZ Europeans. While NZ Europeans showed no difference between males and females on the EPQ-R, Mâori showed variance for Extraversion and Neuroticism. The greater amount of variance for Mâori gender groups was also present on the GPP-I with five traits showing differences for Mâori compared to only three for NZ Europeans. In addition, to the greater number of traits showing variance that were observed between gender groups for Mâori, Mâori were also the only group out of the two that showed consistency among conceptually related traits. This suggests that the differences observed for Mâori are more likely to be representative of real trait differences on the respective traits compared to the gender differences observed for NZ Europeans.

Pairing this result with past research comparing gender differences among ethnic groups provides interesting observations. Dion and Yee (1987) explored the differences between genders for European and Asian ethnic groups. Their findings suggested that Asians had greater gender role differentiation than Europeans. Thus,

indicating a small variance between gender roles for Europeans. The present finding is consistent with Dion and Yee, as Mâori in the present study also appears to show greater gender role differentiation than NZ Europeans. This suggests that the NZ Europeans may have smaller degrees of gender role differentiation than other ethnic groups.

The finding in the present section does show some impact of ethnicity present on gender relevant patterns of personality assessment. Due to the limited research that has explored this area of study, the present findings indicate that this is an area where further investigation maybe beneficial.

7.5. Findings on Personality differences and their Utility in Personnel Selection

In addition to analysing the absence or presence of personality differences among New Zealand ethnic groups, another aim of the present study was also to examine whether differences observed on personality assessments impacts on the procedure of selecting applicants and the outcomes of the selection regime. To explore this issue correlations between personality traits measured by the EPQ-R and the GPP-I and the job offer variable were computed. In addition *t*-test and effect sizes were also produced.

The particular traits shown to be significantly related to selection outcomes in the current study were Psychoticism, Extraversion, and Neuroticism on the EPQ-R. This finding is consistent with previous research administered in Armed Force settings and research on work related behaviours using the EPQ series of assessments (Eugenia, Fuenzalida, Swickert, & Hittner, 2004; Eysenck, 1983; Eysenck & Eysenck, 1985; Jessup & Jessup, 1971; Martin & Kirkcaldy, 1998; Spencer & Spencer, 1966). In addition, this stage of the analysis also shows that Ascendency, Original Thinking, and Vigor on the GPP-I were significantly related to selection outcomes administered by the NZ Army OSB. This finding is slightly less consistent with previous research, as while the Ascendency and Vigor traits are consistent with past research (Gordon,

1978), Original Thinking has not previously been shown to be associated with Armed Forces selection outcomes. Further, while in the present study Responsibility does not appear to be related to selection outcomes past research indicates that this trait predicts selection outcomes (Gordon, 1978). It is worth noting that it is possible that the differences in results for the GPP-I are due to differences in the Job criteria between past studies conducted and the role of Officer in the NZ Army.

The finding that Psychoticism, Extraversion, and Neuroticism on the EPQ-R and Ascendency, Original Thinking, and Vigor on the GPP-I were significantly related to selection outcomes within the context of ethnic differences suggests that, overall, the potential for adverse impact for any ethnic group from both the EPQ-R and GPP-I on employment decisions is relatively small.

Vigor appears to be the only trait of the six traits related to employment decisions to have potential to create adverse effects for New Zealand ethnic groups. This trait appears to show a strong relationship with the job offer variable and also has large differences among ethnic groups with NZ Europeans tending to score higher than all other ethnic groups. The group that appears to be most susceptible to adverse impact is Pasifika, followed by Asians. The risk for Pasifika and Asians to be subjected to adverse impact as a result of Vigor scores is at a magnitude (around .50) that suggests caution should be taken when interpreting the GPP-I for NZ Army Officer selection purposes.

It is worth noting here that, although the Lie scale on the EPQ-R does not appear to have a strong relationship with selection outcomes, due to the nature and purpose of the scale, the large difference reported in the results among ethnic groups indicates some concerns for the NZ Army OSB. In addition, in the OSB selection process the Lie scale is used as a "Red Flag" indicator, which means that under certain circumstance this scale is likely to contribute to the selection outcomes of Officer applicants. The major concern among ethnic groups is for Asians, as this group scored much higher on the Lie scale than NZ Europeans in the present analysis. The concern rises not only because the Asian group scored higher than other ethnic groups but also because the interrelationships between the traits and the Lie scale on the EPQ-R, discussed earlier (cf. section 7.1), suggest that this scale for Asians is not an accurate

measure of "faking good". It is therefore important to consider whether the correct judgement for the interpretation of this scale is to provide special norms for the Asian group. In other words, due to the use of the Lie scale by the NZ Army OSB (i.e. used as a "red flag") it might be helpful to use separate Asian Lie scale norms for this cut-off score decision rule to help reduce the effect of adverse impact.

It should be noted that this interpretation and recommendation for the Lie scale is not likely to be an appropriate procedure to accommodate for the large differences observed among ethnic groups for Vigor. Because Vigor is a scale measuring behavioural tendencies based on trait theory, the strategy of using corrected and separate norms among ethnic groups on the Vigor trait would not be beneficial. The major reason for this is that Vigor, like other personality traits is based on theory suggesting that certain levels of Vigor are associated with certain magnitudes of behaviours incorporated in ones personality. Therefore, it is not recommended that the same strategy used for the correction of the Lie scale be applied for the Vigor trait. Furthermore, because the GPP-I is used as a behavioural indicator rather than a screening "red flag" indicator it is likely that because 59.3% of applicants that undertook both the EPQ-R and GPP-I in the present study were selected by the NZ Army OSB, only very large differences among ethnic groups will translate into adverse impact for certain ethnic groups (Roth et al. 2001). Nevertheless, it is recommended that at the very least, caution with Vigor scores is taken into account when using the GPP-I for selection in the NZ Army.

Overall, the analysis found that personality traits were significantly related to selection outcomes, thus, supporting past research investigating the relationship between personality traits and selection (Barrick & Mount, 1991; Barrick & Mount, 1993; Barrick & Mount, 1996; Barrick, Mount, & Struass, 1993; Hogan & Holland, 2003; Judge, Bono, Erez, Locke, & Thoresen, 2002; Judge & Ilies, 2002; Mount, Barrick, & Struass, 1994; Tett et al. 1991). The differences among ethnic groups in the present study do not appear to suggest the potential for adverse impact when using the GPP-I and EPQ-R for selection purposes in the NZ Army. This finding suggests that overall the two inventories used in the NZ Army OSB meet within ethical principles (New Zealand Psychological Society, 2002) and legislative requirements (Employment Relations Act, 2000) in New Zealand. However, it is suggested that

some caution should be used when interpreting the Lie scale on the EPQ-R for Asian groups and the Vigor trait for Mâori, Asian, and Pasifika groups, as the magnitude of these differences on these traits have potential for adverse impact.

8. Limitations and Future Perspectives

8.1. Limitations

In addition to the limitations outlined at the very beginning of this discussion, the present study has several limitations to the findings that require consideration. Firstly, there are further complications involved with the sample groups used in the present study. Similar to other studies that have analysed the influence of demographic variables on personality assessment the sample size of some of the ethnic groups used in the present study were small. While the NZ European sample was acceptable the ethnic groups of Mâori (N =64), Asian (N = 20), and Pasifika (N = 17) were relatively small, which is possibly a contributing factor to the many non-significant findings. Specifically, the ethnic groups of Asian and Pasifika were very small and close to a level that can generate statistical problems during a statistical analysis. Thus, the ability to generalize the absence or presence of group differences found in the present study to wider New Zealand populations is limited.

The sample groups used in the present study are also limited by an additional issue. The identification of ethnic group for participants in the study was determined by their response on a self-report question asking "What ethnic group do you identify with?" This method is problematic for two reasons. Previous research suggests this response is likely to show variance over time and situation settings. Thus, it is possible that ethnic group status for some participants is subject to change. This is reinforced by the fact that within New Zealand one in ten (10.4%) individuals identify with more than one ethnic group (New Zealand census data 2006). Therefore, within this context the ethnic groups that we have examined may not provide a correct representation of the groups that we are trying to measure in the present analysis.

In addition to sampling issues there is also some level of range restriction imposed upon the present study. Although the present study is not restricted as much as other studies in this area the data accessed was from the second stage of the analysis in the NZ Army OSB selection process. This is an important aspect of the present research, as the NZ Army OSB selected applicants for the second stage of the selection process

upon results obtained from a cognitive ability test, an essay-writing test, a medical screening form, and an interview with a recruitment officer. Therefore it is possible that this may have been a contributing factor to the small amount of differences found between the different ethnic groups in this study.

Due to the lack of available data, the present study was also limited to analysing only four New Zealand ethnic groups within New Zealand. Previous research suggests that it is possible for other ethnic groups not included in the present study to show significantly different results compared to those found in the present study (Eaves et al. 1986).

8.2. Future Research

There are many areas involving the relationship between personality and ethnicity that are in need of further investigation. The sparse number of studies means there is room for many more similar in nature to the current one. Administering such research will allow us to establish patterns and inconsistencies that will provide further knowledge and insight regarding not just the relationship between ethnicity and personality but the influences and origins of different personality traits.

The present study has provided findings, which suggest a general lack of consistency across personality inventories and personality models. This indicates the possibility that ethnic group differences observed on personality assessments may be a result of measurement artifacts rather than real trait differences. Resolving this issue is beyond the scope of the present study due to the design and the characteristics of the data collected. However, it is possible that with the right research design further insight and understanding can be achieved. It is suggested that researchers should further investigate this issue by examining the differences among ethnic groups observed on personality questionnaires, with other behavioural, neurological, and biological outcomes related to specific models and theories of personality. For example, this could be investigated by analysing the consistency of differences among ethnic groups on Extraversion measured by the EPQ-R (or EPQ) with evoked-potential records theoretically related to different levels of Extraversion (Eysenck, 1947). Another possibility would be to compare the ethnic differences on self-report

personality questionnaires with differences in observer-ratings (i.e. analysing the consistency of phenotype and genotype traits). Such research could be invaluable for furthering understanding of personality differences and to enable personality research to disentangle the complexities it faces when confronting ethnic group differences and adverse impact issues.

The findings of the present study were restricted by the level of measurement, as they were obtained from only the highest facet level for both the EPQ-R and GPP-I. An enchanting avenue for further research could be to investigate the similarities and inconsistencies of ethnic group differences across different facet levels of trait measurement. For example, for personnel selection, examination of ethnic differences could be compared on a personality inventory between the global traits and the sub facet level traits. In addition, it may be of further interest to examine the presence or absence of differences among ethnic groups at the item level of measurement. At this level very specific differences could be examined and investigated. Further, exploring the ethnic effects on personality assessment does not necessarily have to stop at the item level of measurement. To fully understand what ethnic test differences mean in terms of adverse impact, fairness and bias, research may also need to examine this issue at the profile interpretation level, prediction level, and decision level.

Another strategy, similar to that prescribed above, would be the relationship between personality assessment among ethnic groups and corresponding job performance ratings or scores given. Johnson, Carter, Davison and Oliver (2001) suggest that this issue can be examined by analysing whether the regression equation between personality and performance differs for different ethnic groups. Such research could be administered using the same design as earlier studies investigating the relationship between personality and job performance (e.g. Barrick & Mount, 1991, 1996; Barrick, Mount, & Struass, 1993; Tett et al. 1991) by simply including the additional variable of ethnic group.

While examining the issue of measurement equivalence in selection settings is very important for HR specialists and the continued use of personality assessment in industrial settings, it is perhaps even more important for further research to examine this concept within the context of the theory behind models of personality. Moreover,

it may be of interest to investigate this issue by comparing personality differences among ethnic groups with other theoretically related outcomes. For example, PEN differences among ethnic groups could be compared with different behaviours, personality disorders (e.g. anxiety, depression), correlations with psychological distress, neurological activity, evoke potentials etc. The importance of such research cannot be overstated, as this type of analysis could provide further insight into the theoretical rationale for ethnic group differences and perhaps even provide evidence for refinements of fundamental theories of personality. Further, such revelations in personality research could be very important to the validity of personality assessments for employment settings as they could provide further confirmation of the relationship between personality assessment scores and job performance.

In addition to mean trait differences among ethnic groups the present study also investigated the influence of ethnicity on the relationship between gender and personality. The present study has highlighted the importance of this area, as it found observable differences between Mâori and NZ Europeans for gender group differences. This analysis suggested that different ethnic groups might have different role differentiation levels. Like direct difference among ethnic groups on personality traits, this type of analysis has not been extensively examined. Thus, further analysis in this area is imperative for the same reasons as given for direct ethnic differences (i.e. personality theory, valid measurement).

9. Conclusion

To conclude, this study has presented the findings of one of the first independent evaluations on the potential impact of demographic factors such as ethnicity on personality traits, such as those measured by the EPQ-R and the GPP-I. The current study had six research aim:

(vii) Analyse New Zealand data to find potential relationships between personality and ethnicity. This included comparing ethnic groups along relevant personality variables (e.g. by means of two-tailed *t* tests utilising the means as well as effect size differences and standard deviations of ethnicity defined subsets among the participants), as well as an analysis of the impact of ethnic status on gender- and age- relevant patterns in personality assessment.

(viii) Provide further insight as to whether statistical comparisons between ethnic groups reflect "real" similarities and differences.

(ix) Provide information for New Zealand ethnic groups on other models of personality besides the FFM that are currently used in personality research and personnel selection.

(x) Check whether any relationships between personality and ethnicity can be shown to be pervasive, irrespective of choice of a (systemic) model.

(xi) Examine whether differences observed on personality assessment scores significantly impacts on employee selection outcomes.

(xii) Provide further data on the relationship between personality and ethnicity for New Zealand sample groups.

The results in the present study indicated that ethnicity did have some impact on the personality assessment scores used in the study. The relationship appeared to be of a

complex nature, as the same pattern of identity and difference was not found consistently across ethnic groups on each personality trait. Rather exactly which trait(s) that were shown to differ between any two ethnic groups was largely dependent on the specific ethnic groups being compared. Apart from the Lie scale, Personal Relations, and Vigor, which reach moderate levels, most of the differences were quite small by Cohen's standards (1988). Both the *d* statistic and discriminant analysis were consistent with these small differences observed, as only one significant discriminant function was produced in the present analysis. Moreover, there were some indications that for the Pasifika group GPP-I trait scores compared to other ethnic groups were substantially different.

The majority of the differences found on the EPQ-R and the GPP-I were not consistent across the two inventories and also with similar ethnic groups analysed using the FFM (i.e. Packman et al. 2005). The inconsistency found in the study questions the validity of these differences. In other words, the present findings raise the possibility that the differences observed among the ethnic groups are not a result of real differences but rather a result of artifacts created by the personality measures. It is suggested that future research investigate this issue further.

The present study did not find any significant differences for the variance on personality traits among ethnic groups. This result provides indication that ethnicity does not influence the variance of personality traits. However, due to the small amount of research conducted on this issue it is suggested that future research continue to investigate this. While initial analysis suggested that ethnicity did not influence the relationship between age and gender on personality assessment further examination suggested that the relationship between gender and personality was impacted by ethnicity. The pattern of this impact suggested that NZ Europeans have smaller degrees of gender role differentiation than other ethnic groups. This pattern was also consistent with past research comparing Europeans to other ethnic groups (Dion & Yee, 2001). These findings highlight the need for further research in this area as the research conducted so far on this issue is sparse.

Despite the differences among ethnic groups on the personality traits, the present study did find reassuring results for New Zealand psychologists, HR specialists, and employers. Traits that showed a relationship with employment decisions mostly only had small differences among ethnic groups. However, there were some traits identified that in certain circumstances should be interpreted with caution. In addition, for the NZ Army selection process the findings in the present study indicate that the Lie scale on the EPQ-R and the Vigor trait on the GPP-I may need to be interpreted with caution for Asian and Pasifika groups. Careful attention needs to be taken particularly for the Lie scale on the EPQ-R, as the relationship it has with Neuroticism in the EPQ-R suggests that it is not an accurate measure for faking good for the Asian ethnic group.

Overall, the findings of the present study indicate that ethnicity has some impact on personality assessment. However, the magnitude of this impact appears to be relatively small compared to other forms of assessment in personnel selection (e.g. Cognitive Ability). However, due to the design and methodological shortcomings restricting the present study, it is suggested that further research continue to explore this issue. Further research in this area is important as such studies may not only provide greater understanding of the potential of adverse impact in personality assessment, but also provide us with greater understanding of the fundamental theory of personality.

Reference

Acton, G. S. (2003). Measurement of impulsivity in a hierarchical model of personality traits: Implications for substance use. *Substance Use & Misuse, 38,* 67-83.

Adelson, J. (1969). Personality. *Annual Review of Psychology*, *20*, 136-252.

Agresti, A. & Finlay, B. (1997). *Statistical Methods for the Social Sciences* (3rd ed.). Upper Saddle River, New Jersey: Prentice Hall.

Allan, J. (1998). Personality Assessment with American Indians and Alaska Native: Instrument Considerations and Service Delivery Style. *Journal of personality Assessment*, *70*, 17-41.

Allan, J. (2001). *Review of the Measurement of Ethnicity-Classification and Issues.* Classifications and Standards Section. Statistics New Zealand.

Alker, H. A. (1972). Is personality situationally specific or intrapsychically consistent? *Journal of Personality*, *40,* 1-16.

Allport, G.W. (1937). *Personality: A psychological interpretation.* New York: Holt, Rinehart & Winston.

Allport, G.W. (1961). *Pattern and growth in personality*. New York: Holt, Rinehart & Winston.

Allport, G.W., & Odbert, H.S. (1936). Trait names: a psycho-lexical study. *Psychological Monographs*, *47*, 211.

Arima, J. K. (1967). Differentiation of soldier reactions to severe environmental stress by MMPI and other variables. *Educational and Psychological Measurement*, 27, 617-630.

Barrett, P. T., & Eysenck, S. B. G. (1984). The assessment of Personality across 25 countries. *Personality and Individual Differences, 5*, 615-632.

Barrett, P. T., Petrides, K. V., Eysenck, S. B. G., & Eysenck, H. J. (1998). The Eysenck Personality Questionnaire: An examination of the factorial similarity of P, E, N, and L across 34 countries. *Personality and Individual Differences*, *25*, 805–819.

Barrick, M. R., & Mount, M. K. (1991). The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. *Personnel Psychology, 44*(1), 1-26.

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, *81*, 261-272.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and Performance at the Beginning of the New Millennium: What Do We Know and Where Do We Go Next? *International Journal of Selection and Assessment, 9* (1&2), 9-30.

Barrick, M. R., Mount, M. K., & Strauss, J. P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. *Journal of Applied Psychology*, *78*, 715-722.

Barrett, P.T., Petrides, K.V., Eysenck, S.B.G., & Eysenck, H.J. (1998) The Eysenck Personality Questionnaire: an examination of the factorial similarity of P, E, N, and L across 34 countries. *Personality and Individual Differences, 25*(5), 805-819.

Bartram, D., & Dale, H. C. A. (1982). The Eysenck Personality Inventory as a selection test for military pilots. *Journal of Occupational Psychology*, *55*, 287-296.

Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review, 81*, 506-520.

Ben-Porath, Y.S. & Waller, N.G. (1992a). Five big issues in clinical personality assessment: A rejoinder to Costa and McCrae. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *4*, 23-25.

Ben-Porath, Y. S., & Waller, N. G. (1992b). "Normal" personality inventories in clinical assessment: General requirements and the potential for the using the NEO Personality Inventory. *Psychological Assessment, 4,* 14-19.

Bishop, D. V. M. (1977). The P scale and psychosis. *Journal of Abnormal Psychology*, 86, 127-134.

Black, J. (1998, September). *Does a job applicant's personality contribute to a successful job application?* Paper presented at the annual conference of the New Zealand Psychological Society, Wellington, NZ.

Black, J. (2000). Personality testing and police selection: Utility of the 'Big Five'. *New Zealand Journal of Psychology*, *29*, 2-9.

Block, J. (1977). The Eysencks and Psychoticism. Journal of Abnormal Psychology, *86*, 653-654.

Block, J. (1995). A contrarian view of the Five-Factor approach to personality description. *Psychological Bulletin, 117*(2), 187-215.

Block J (2001) Millennial contrarianism: The Five Factor approach to personality description 5 years later. *Journal of Research in Personality*, *35*, 98-107.

Buss, A. H. (1989). Personality as traits. *American Psychologist*, *44*, 1378-1388.

Butcher, J. N., Dahlstrom, W.G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Manual for the restandardized Minnesota Multiphasic Personality Inventory: MMPI-2. An administrative and interpretive guide*. Minneapolis, MN: University of Minnesota Press.

Bouchard, T. J. (2004). Genetic influence on human psychological traits: A survey. *Current Directions in Psychological Science, 13*, 148–151.

Bowers, K. S. (1973). Situationism in psychology: An analysis and a critique. *Psychological Review*, *80*(5), 307-336.

Braun, J. R. (1963). Fakability of the Gordon Personal Inventory: Replication and extension. *Journal of Psychology, 58*, 441-444.

Braun, J. R., & Gomez, B. J. (1966). Effects of faking instructions on the Eysenck Personality Inventory. *Psychological Reports, 19,* 388-90.

Bronfenbrenner, U. (1953). Personality. *Annual Review of Psychology*, *4*,157-82.

Brody, N. (1988). *Personality: In search of individuality*. New York: Acedemic.

Brislin. R. W. (1983). "Cross-Cultural Research in Psychology." *Annual Review of Psychology, 34*, 363-400.

Bryson, J., & Hosken, C. (2005). What does it mean to be a culturally competent I/O psychologist in New Zealand? *New Zealand Journal of Psychology, 34*(2), 69(68).

Buss, A. H. (1989). *Personality: Evolutionary heritage and human distinctiveness*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Carlson, R. (1971) "Where is the person in personality research?" *Psychological Bulletin*. 75. 203- 219.

Carmon, N. (Ed.). (1996). *Immigration and Integration in Post-Industrial Societies*: *Theoretical analysis and policy related research.* Basingstoke, UK: MacMillian Press Ltd.

Cervone, D., & Pervin, L. A. (2008). *Personality: Theory and Research* (10th ed.). New York: John Wiley & Sons.

Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika, 12*, 197-220.

Cattell, R. B. (1948). The primary personality factors in women compared with those in men. *British Journal of Psychology, 1*, 114-130.

Chan, D. (1997). Racial Subgroup Differences in Predictive Validity Perceptions on Personality and Cognitive Ability Tests. *Journal of Applied Psychology, 82*(2), 311-320.

Cheung, F. M., & Leung, K. (1998). Indigenous personality measures: Chinese examples. *Journal of Cross-Cultural Psychology, 29*, 216-233.

Cheung, F. M., Leung, K., Zhang, J. X., Sun, H. F., Gan, Y. Q., Song W. Z., & Xie, D. (2001). Indigenous Chinese personality construct: Is the Five Factor Model complete? *Journal of Cross-Cultural Psychology, 32,* 407-433.

Church, A. T. (2000). Culture and personality: Towards an integrated cultural trait psychology. *Journal of Personality*, 68, 651-703.

Church, A.T. (2001). Personality measurement in cross-cultural perspective. *Journal of Personality, 69*(6), 979-1007.

Church, A., & Lonner. W. (1998). The cross-cultural perspective in the study of personality: Rationale and current research. *Journal of Cross-Cultural Psychology, 29*, 32-62.

Cliff, N. (1983). Some cautions concerning the application of causal modelling methods. *Multivariate Behavioural Research, 18*, 115-126.

Cohen, J. (1988). *Standard Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Collins, J. M. & Gleaves, D. H. (1998). Race, job applicants, and the Five Factor model of Personality: Implications for black psychology, industrial /organizational psychology, and the Five-Factor theory. *Journal of Applied Psychology, 83*, 531-544.

Cook, M., & Cripps, B. (2005). *Psychological assessment in the workplace: A manager's guide*. Chichester, England: Hoboken, NJ: Wiley.

Coolican, H. (2006) *Introduction to Research Methods in Psychology* (3rd ed.). London: Hodder & Stoughton.

Costa P.T., Jr. (1991). Clinical use of the five-factor model: An introduction. Special Series: Clinical use of the five-factor model of personality. *Journal of Personality and Assessment, 57*, 393-398.

Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO personality inventory manual*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T, Jr, & McCrae, R. R (1988). From cataloge to classification: Murray's needs and the five-factor model. *Journal of Personality and Social Psychology*, *55*, 258-265.

Costa, P. T., & Widiger, T. A. (1994). *Personality Disorders and the Five-Factor Model of Personality.* Washington, DC: American Psychological Association.

Cox-Fuenzalida, L. E., Swickert, R. J., & Hittner, J. B. (2004). Effect of Neuroticism and workload history on performance. *Personality and Individual Differences*, *36*, 447-456.

Cowles, M., Darling, M., & Skanes, A. (1992). Some characteristics of the similated self. *Personality and Individual Differences, 13,* 501-510.

Cramer, D. (2003). *Advanced Quantitative Data Analysis*. Maidenhead (Berkshire, United Kingdom): Open University Press.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist, 30,* 671–84.

Cronbach, L. J., & Meehl, P. C. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302.

Dawes, R. M. (1994). *House of Cards: Psychology and Psychotherapy Built on Myth*. New York: The Free Press.

De Dreu, C. K. W., & West, M. A. (2001). Minority Dissent and Team Innovation: The Importance of Participation in Decision Making. *Journal of Applied Psychology, 86*(6), 1191-1201.

Digman, J. M. (1990). Personality structure: Emergence of the Five Factor Model. *Annual Review of Psychology, 41,* 417-440.

Dion, K. L., & Yee, P. H. N. (1987). Ethnicity and personality in a Canadian context. *Journal of Social Psychology, 127*, 175-182.

Draycott, S. G., & Klein, P. (1995). The big three or the big five – the EPQ-R vs the NEO-PI: A research note, replication and elaboration. *Personality and Individual Differences, 6,* 801-804.

Driskell, J. E., Hogan, J., Salas, E., Hoskin, B. (1994). Cognitive and personality

predictors of training performance. *Military Psychology, 6*(1), 31-46.

Eaves, L. J., Eysenck, H. J., & Martin, N. G. (1989) *Genes, Culture and Personality*. London: Academic.

Ehud, B., Elishera, B., & Zeev, K. (2006). Soldiers who kill themselves: The contribution of dispositional and situational factors. *Archives of Suicide Research, 10*(1), 29-43.

Employment Relations Act (2004). (accessed from www January 12th 2008: http://union.org.nz/policy/employment-relations-act-2004-consolidated).

Epstein, S. (1979). The stability of behavior: on predicting most of the people much of the time. *Journal of Personality and Social Psychology. 37*, 1097-1126

Epstein, S. (1984). *Controversial issues in emotion theory*. In P. Shaver (Ed.), Annual review of research in personality and social psychology (pp. 64-87). Beverly Hills, CA: Sage Publications.

Evers, A., Anderson, N., & Voskuijl, O. F. (Eds.) (2005). *The Blackwell Handbook of Personnel Selection*. Oxford, England: Blackwell.

Eysenck, H. J. (1947). *Dimensions of Personality*. London, U.K: Routledge & Kegan Paul.

Eysenck, H. J. (1967). *The Biological Basis of Personality.* Springfield, IL: Charles C. Thomas.

Eysenck, H.J. (1977) *Crime and personality*, 3rd ed. Routledge & Kegan Paul Pub.

Eysenck, H. J. (1982). *Personality, genetics, and behavior.* New York: Praeger.

Eysenck, H. J. (1990a). Biological dimensions of personality. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 244-276). New York: Guilford.

Eysenck, H. J. (1990b). Genetic and environmental contributions to individual differences: The three major dimensions of personality. *Journal of Personality, 58,* 245-261.

Eysenck, H. J. (1991). Dimensions of personality: 16, 5, or 3?--Criteria for a taxonomic paradigm. *Personality and Individual Differences, 12,* 773-790.

Eysenck, H.J. (1992a). A reply to Costa and McCrae. P or A and C - the role of theory. *Personality and Individual Differences*, *13*, 867-868.

Eysenck, H.J. (1992b). Four ways five factors are not basic. *Personality and Individual Differences*, *13*, 667-673.

Eysenck, M.W. (1983). Anxiety and individual differences. In R. Hockey (Ed.), *Stress and Fatigue in Human Performance*. Chichester: Wiley

Eysenck, H. J., & Eysenck, S.B.G. (1970). A factor-analytic study of the Lie scale of the Junior Eysenck Personality Inventory. *Personality, 1*, 3-10.

Eysenck, H. J., & Eysenck, S.B.G. (1975). Manual of the Eysenck Personality *Questionnaire (adult and junior)*. London: Hodder and Stoughton.

Eysenck, H.J., & Eysenck, S.B.G. (1976). *Psychoticism as a Dimension of Personality*. London: Hodder and Stoughton.

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and Individual Differences: A natural science approach.* New York: Plenum.

Eysenck, H. J., & Eysenck, S. B. G. (1991). *Eysenck Personality Scales (EPS Adult).* London: Hodder & Stoughton.

Eysenck, S. B. G., Barrett, P. T., & Barnes, G. E. (1992). A cross-cultural study of personality: Canada and England. *Personality and Individual Differences, 14*(1), 1-10.

Eysenck, S. B. G., Nias, D. K. B., & Eysenck, H. J. (1971). The interpretation of children's Lie scale scores. *British Journal of Educational Psychology, 41*, 23-31.

Eysenck, S. B. G., Syed, I. A., & Eysenck, H. J. (1965). Desirability response set in children. *British Journal of Educational Psychology, 36,* 87-90.

Eysenck, S.B.G., & Chan, J. (1982). A comparative study of personality in adults and children: Hong Kong vs. England. *Personality and Individual Differences, 3*, 153–160.

Farley, F. H., & Goh, D. S. (1976). PENmanship: faking the P-E-N. *British Journal of Social and Clinical Psychology, 15,* 139–148.

Fisk, D. W. (1974). The limits of the conventional science of personality traits. . *Journal of Personality*, *42,* 1-11.

Frankfort-Nachmias, C., & Nachmias, D. (1996). Sampling and Sample Designs. In Ch. Frankfort-Nachmias & D. Nachmias, *Research Methods in the Social Sciences* (5th ed.) (pp. 177-202). New York: St. Martin Press.

Frei, R., & McDaniel, M. A. (1998). The validity of customer service orientation measures in employee selection: A comprehensive review and meta-analysis. *Human Performance, 11*, 1-27.

Fuenzalida, L.E., Swickert, R.J., & Hittner, J.B. (2005). Effect of Neuroticism and workload history on performance. *Personality and Individual Differences, 36*, 447-456.

Furnham, A. (1981). Personality and activity preference. *British Journal of Social Psychology, 20*, 57-68.

Futuyma, D. J. (1979). *Evolutionary Biology.* 1st ed. Sinauer Associates, Sunderland, Massachusetts.

Galton, F. (1869). *Hereditary Genius*. London: Macmillan.

Galton, F. (1884). Measurement of character. *Fortnightly Review, 36,* 179-185.

Giles, D. C. (2002). *Advanced Research Methods in Psychology*. Hove (Sussex, United Kingdom): Routledge.

Glass, G. V., & Hopkins, K. D. (1996). *Statistical Methods in Education & Psychology* (3rd ed.). Boston: Allyn & Bacon.

Goeters, K., Timmerman, B., & Maschke, P. (1993). The construction of personality questionnaires for selection of aviation personnel. *International Journal of Aviation Research, 3,* 123-141.

Goffin, R. D., Rothstein, M. G., & Johnston, N. (1996). Personality testing and the assessment center: Incremental validity for managerial selection. *Journal of Applied Psychology, 81*, 746-756.

Goldberg, L. R. (1982). From Ace to Zombie: Some explorations in the language of personality. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in Personality Assessment* (Vol. 1, pp. 203-234). Hillsdale, NJ: Erlbaum.

Goldberg, L. R. (1990). An alternative "Description of Personality": The Big-Five Factor Structure. *Journal of Personality and Social Psychology,* 59, 1216-1229.

Goldberg, L. R. (1993). The Structure of Phenotypic Personality Traits. *American Psychologist, 48*(1), 26-34.

Goldberg, L. R., & Kilkowski, J. M. (1985). The prediction of semantic consistency in self-descriptions: Characteristics of persons and of terms that affect the consistency of responses to synonym and antonym pairs. *Journal of Personality and Social Psychology*, *48*, 82-98.

Goldberg, L. R., Sweeney, D., Merenda, P. F., & Hughes, J. E. (1998). Demographic variables and personality: the effects of gender, age, education, and

ethnic/racial status on self-descriptions of personality attributes. *Personality and Individual Differences, 24*(3), 393-403.

Goodstein, L. D., & Lanyon, R. I. (1999). Applications of Personality Assessment to the Workplace: A Review. *Journal of Business and Psychology, 13*(3), 291-322.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Goyne, A. (2006). Personality Testing in the Australian Defence Force: An Historical Review. Psychology Research and Technology Group Australian Regular Army. (Accessed on February 18th 2008:http://209.85.141.104 /search?q=cache:_kpuL7V8s88J:www.internationalmtaorg/Documents/2006/006012P.pdf+Goyne,+2006+personality+job+performance&hl=en&ct=clnkcd=2&gl=nz).

Guanzon-Lapena, M. A., Church, A. T., Carlota, A. J., & Katigbak, M. S. (1998). Indigenous personality measures: Philippine examples. *Journal of Cross-Cultural Psychology, 29*(1), 249-270.

Guenole, N., & Chernyshenko, O. S. (2005). The suitability of Goldberg's Big Five IPIP personality markers in New Zealand: a dimensionality, bias, and criterion validity evaluation. *New Zealand Journal of Psychology, 34*(2), 86(11).

Guion, R. M. (1998). Gordon Personal Profile-Inventory (Revised). In J. C. Impara & B. S. Plake (Eds.). *The Thirteenth Mental Measurements Yearbook*. Buros Institute.

Guion, R. M., & Gottier, R. F. (1965). Validity of Personality Measures in Personnel Selection. *Personnel Psychology*. 18, 135-164.

Gustavsson, J. P., Weinryb, R. M., Goransson, S., Pedersen, N. L., & Asberg, M (1997). Stability and predictive ability of personality traits across 9 years.. *Personality and Individual Differences, 22,* 783-791.

Hansen, M. H., Hurwitz, W. N., & Madow, W. G. (1993). *Sample survey methods and theory*. Vols. I. & II. New York: Wiley.

Hartmann, E., Sunde, T., Kristensen, W., & Martinussen, M. (2003). Psychological measures as predictors of military training performance. *Journal of Personality Assessment, 80*, 87-98.

Haslam, N. (2007). *Introduction to Personality and Intelligence*. London: Sage Publications.

Hess, A. K. (1998). Gordon Personal Profile-Inventory (Revised). In J. C. Impara & B. S. Plake (Eds.). *The Thirteenth Mental Measurements Yearbook*. Buros Institute.

Heuchert, J. W. P., Parker, W. D., Stumpf, H., & Myburgh, C. P. H. (2000). The Five-Factor Model of Personality in South African College Students. *American Behavioral Scientist*, 44(1), 112-125.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied Statistics for the Behavioral Sciences*. Boston: Houghton Mifflin Co.

Hoffmann, T., Dana, R. H., & Bolton, B. (1985). Measured acculturation and MMPI-168 performance of Native American adults. *Journal of Cross-Cultural Psychology*, *16*(2), 243-256.

Hogan, J., & Holland, B. (2003). Using Theory to Evaluate Personality and Job-Performance Relations: A Socioanalytic Perspective. *Journal of Applied Psychology, 88*(1), 100-112.

Hogan, R., Johnson, J., & Briggs, S. (1997). *Handbook of Personality Psychology* (NY: Academic Press).

Hogan, R. & Jones.W. (1985). *Perspectives in personality: Theory, measurement, and interpersonal dynamics*. Greenwich, CT: JAI Press.

Hough. L. M. (1992). The "Big Five" personality variables-construct confusion: Description versus prediction. *Human Performance, 5,* 139-155.

Hough, L. M. (Ed.). (1998a). *Personality at work: Issues and evidence*. Hillsdale, NJ: Erlbaum Associates.

Hough, L. M. (1998b). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance, 11*(2/3), 209.

Hough, L., & Furnham, A. (2003). Importance and use of personality measures in work settings. In R.D. Hackett (ED.), Comprehensive Handbook of Psychology Vol. 12: W.Borman, D. Ilgen, & R.Klimoski (Eds.), Industrial/Organizational Psychology. New York: Wiley & Sons. pp. 131-169.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

International Labour Organization (ILO). (2004). Combating discrimination against migrant and minority workers. (Accessed on November 5th, 2007: http://www.ilo.org/public/english/protection/migrant/projects/discrim/).

Ivkovic,V., Vitart,V., Rudan, I., Janicijevic, B., Smolej-Narancic, N., Skaric-Juric, T., Barbalic, M, Polasek, O, Kolcic, I., Biloglav, Z., Visscher, P. M., Hayward, C., Hastie, N. D., Anderson, N., Campbell, H., Wright, A. F., Rudan, P., & Deary, I. J. (2007). The Eysenck personality factors: psychometric structure, reliability, heritability and phenotypic and genetic correlations with psychological distress in an isolated Croatian population. *Personality and Individual Differences*, *42*, 123-133.

Jackson, C. J. & Francis, L. J. (1999). Interpreting the correlation between Neuroticism and Lie scale scores. *Personality and Individual Differences, 26*, 59-63.

Jessup, G., & Jessup, H. *(*1971*).* Validity of the Eysenck Personality Inventory in *pilot selection. Occupational Psychology, 45*, 111-123.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66-100). New York: Guilford.

John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed.)(pp. 102-138). New York: Guilford Press

Judge, T. A, Bono, J. E., Erez, A., & Locke, E. A. (2005) Core self-evaluations and job and life satisfaction: The role of self-concordance and goal attainment. *Journal of Applied Psychology, 90,* 257-268.

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, Neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology,* 83, 693-710.

Judge, T. A., & Ilies, R. (2002). Relationship of personality to performance motivation: A meta-analysis. *Journal of Applied Psychology, 87*, 797-807.

Jung, C.G. (1921). *Psychological Types*. Princeton, NJ: Princeton University Press.

Kalton, G. (1983). Introduction to Survey Sampling. Newbury Heights, CA: Sage.

Keller, M. C., Coventry, W. L., Heath, A. C., & Martin, N. G. (2005). Widespread evidence for non-additive genetic variation in Cloninger's and Eysenck's personality dimensions using a twin plus sibling design. *Behavior Genetics, 35*, 707–721.

Kilkowski, J. M. (1976). An empirical investigation of the consistency of self-ratings. Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

Kirkcaldy, B. D., & Martin, T. (2000). Job stress and satisfaction among nurses. *Organizational Behavior and Human Performance*, *32*, 160–177.

Kline, P. (1993). *The handbook of psychological testing*. London: Routledge.

Klecka, W. R. (1980). *Discriminant Analysis*. Beverly Hills, CA: Sage.

Knapstein, J, W. (1970). A cross-cultural study of certain personality features of tuberculous alcoholic patients. *Dissertation Abstracts International*, *31*, 6260B.

Konstabel, K., Realo, A., & Kallasmaa, T. (2002). Exploring the sources of variations in the structure of personality traits across cultures. In R. R. McCrae & J. Allik (Eds.), *The Five-Factor Model across cultures* (pp. 29-52). New York: Kluwer Academic/Plenum Publishers.

Lachenbruch, P. A. (1975). *Discriminant Analysis*. New York: Hafner Press /London: Collier Macmillan.

Lake, R. E., Eaves, L. J., Maes, H. H., Heath, A. C., & Martin, N. G. (2000). Further evidence against the environmental transmission of individual differences in Neuroticism from a collaborative study of 45,850 twins and relatives on two continents. *Behavior Genetics, 30*, 223–233.

Lazerwitz, B. (1968). Sampling Theory and Procedures. In H. H. Blalock & A. Blalock (Eds.), *Methodology in Social Research* (pp. 278-328). New York: McGraw-Hill.

Levin, J., & Montag, I. (1987). The effect of testing instructions for handling social desirability on the Eysenck Personality Questionnaire. *Personality and Individual Differences, 8*, 163-167.

Lilienfield S. O., Wood, J.M., & Garb, H. (*2000*). The scientific status of projective. techniques. *Psychological Science in the Public Interest*, *2*, 27-66.

Livesley, W. J., & Bromley, D. B. (1973). *Person Perception in Childhood and Adolescence*. London: Wiley.

Long, N., Chamberlain, K., & Vincent, C. (1992). The health and mental health of New Zealand Vietnam war veterans with posttraumatic stress disorder. *The New Zealand Medical Journal, 105,* 417-419.

Lujan, S., & Zapata, J. T. (1983). Personality differences among Mexican-American college freshman. *Journal of College Student Personnel.* (23), 105-111.

Lung, F., Lee, F., & Shu, B. (2002). The premorbid personality in military students with adjustment disorder. *Military Psychology, 18*(1), 77-88.

Lykken, D.T. (1971). Multiple factor analysis and personality research. *Journal of. Research in Personality, 5*, 161-170.

Lynn, R. (1991) *Essays in Honour of Hans Eysenck*. London: Pergamon Press.

McAdams, D. P. (1992). The five-factor model in personality: A critical appraisal. *Journal of Personality*, *60*, 329-361.

McAdams, D. P. (Ed.). (1997). *A conceptual history of personality psychology*. San Diego: Academic Press.

McCrae R. R. E. (1992). The five-factor model: Issues and applications. *Journal of Personality* 60, 329-361.

McCrae, R. R. (Ed.). (2000). Personality Traits and Culture: New Perspectives on Some Classic Issues. *American Behavioral Scientist, 44*(1), 112-114.

McCrae, R. R. (2001a). Traits through time. *Psychological Inquiry*, *12*, 85-87.

McCrae, R. R. (2001b). Trait psychology and culture: Exploring intercultural comparisons. *Journal of Personality*, 69: 819-846.

McCrae, R. R. (2002). NEO-PI-R data from 36 cultures: Further intercultural comparisons. In R. R. McCrae, & J. Allik. (Eds.): *The Five-Factor Model across cultures*. New York: Kluwer Academic/Plenum Publishers.

McCrae, R. R. (2004). Human nature and culture: a trait perspective. *Journal of Research in Personality. 38*(1), 3-14.

McCrae, R. R., & Allik, J. (Eds.). (2002). *The Five-Factor Model* of Personality *across Cultures*. New York: Kluwer Academic/Plenum Publishers.

McCrae, R. R., & Costa, P. T., Jr. (1990). *Personality in Adulthood.* New York: Guilford.

McCrae, R.R., & Costa, P.T., Jr. (1997). Personality trait structure as a human universal. *American Psychologist, 52*, 509-516.

McCrae, R. R., & Costa, P. T., Jr. (1999). *A five-factor theory of personality*. In L. A. Pervin & O. P. John (Eds.), Handbook of personality: Theory and research (2nd ed., pp. 139–153). New York: Guilford Press.

McCrae, R. R., Costa, P. T., Jr., Del Pilar, G. H., Rolland, J. P., & Parker, W. D. (1998). Cross-cultural assessment of the five-factor model: The Revised NEO Personality Inventory. *Journal of Cross-Cultural Psychology,* 29, 171-188.

McCrae, R. R., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In L. A. Pervin & OP John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 139-153). New York: Guilford.

McCrae, R. R., & Terracciano, A. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology. 88*(1), 547-561.

McClelland, D. C. (1951). *Personality.* New York: Dryden.

McCulloch, S. (1993). Recent trends in international assessment. *International* Journal *of Selection and Assessment*, *1*, 59-61.

McDougall, W. (1932). *Introduction to Social Psychology*. Scribners: New York.

McManus, M. A., & Kelly, M. L. (1999). Personality measures and biodata: Evidence regarding their incremental predictive value in the life insurance industry. *Personnel Psychology, 52,* 137–148.

MacKinnon D.W. (1951). Personality. *Annual Review of Psychology*, *2*, 113-136.

Maddi, S. R. (1984). Personology for the 1980's. In R. A. Zucker, J. Aronoff and R. I. Rabin (Eds.), *Personality and the prediction of behavior*. New York: Academic Press.

Martin, T., & Kirkcaldy, B. F. (1998). Gender differences on the EPQ-R and attitudes to work. *Personality and Individual Differences, 24,* 1-5.

Martini, P., Mazzotti, E. and Setaro, S., (1996). Factor structure and psychometric features of the Italian version for the EPQ-R. *Personality and Individual Differences. 21*(6), 877-882.

Matthews, M., Duckworth, A.L., & Kelly, D.R. (2007, May). *Personality Predictors of Military Leadership Performance among West Point Cadets*. Poster session presented at the annual meetingof the American Psychological Society, Washington, D.C.

Medland, F. F., Green, C. G., & Marder, M. (1965). *Psychological factors in the selection of Special Forces officers in report of the eleventh annual human factors research and development conference.* Washington, D. C.: U. S. Army, Office of the Chief of Research and Development.

Meehl, P. E. (1956). Wanted: A good cookbook. *American Psychologist*, *11*, 263-272.

Meehl, P. E. (1992). Factor and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality. 60*, 175-215.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18*, 5–11.

Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749

Mershon, B.*, &* Gorsuch*,* R. L. (1989). Number of factors in the personality sphere: Does increase in factors increase predictability of real life criteria? *Journal of Personality Social Psychology, 55,* 675-680.

Michaelis, W., & Eysenck, H. J. (1971). The determination of personality inventory factor patterns and intercorrelations by changes in real-life motivation. *Journal of Genetic Psychology*, *118,* 223-234.

Middeldorp, C. M., Cath, D. C, van Dyck, R., & Boomsma, D. I. (2005). The co-morbidity of anxiety and depression in the perspective of genetic epidemiology. A review of twin and family studies. *Psychological Medicine*. *35*(5), 611-624.

Miles, R.E. (1989) Adapting to technology and competition: A new industrial relations system for the 21st century. *California Management Review, 31*(1), 9-28.

Miller, R. E., & Ceager, J. A. (1960). *Predicting achievement of cadets in their first year at the Air Force Academy, Class of 1962* (Tech. Note 60-259). Lackland Air Force Base, Tax: Personnel Laboratory, Wright Air Development Division, Air Research and Development Command.

Mischel, W. (1968). *Personality and Assessment*. New York: Wiley.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, *80*, 252-283.

Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the Big Five personality dimensions. *Journal of Applied Psychology*, *79*, 272-280.

Mudge, B. R. (1956). *A study of the Relationship between the Gordon Personal Profile and Academic Achievement*. Unpublished master's thesis, Boston University.

Nakagawa, S. (2004). A farewell to Bonferroni: The problems of low statistical power & publications bias. *Behavioural Ecology*, *15*(6), 1044-1045.

Nemeth, C. J. (1986). The differential contributions of majority and minority influence. *Psychological Review, 93*, 23-32.

Nemeth, C., & Kwan, J. (1985). Originality of word associations as a function of majority vs. minority influence processes. *Social Psychology Quarterly, 48*, 277-282.

Nemeth, C. J., & Kwan, J. (1987). Minority influence, divergent thinking and the detection of correct solutions. *Journal of Applied Social Psychology, 17*, 786-797.

Nemeth, C., & Wachtler, J. (1983). Creative problem solving as a result of majority vs. minority influence. *European Journal of Social Psychology, 13*, 45-55.

New Zealand Census. (2006). New Zealand 2006 Census. (Accessed on November 21st 2008: http://www.stats.govt.nz/census/census-outputs/default.htm).

New Zealand Immigration Service. (2003). *Skilled Immigrant: Labour market experience.* Wellington: Department of Labour.

New Zealand Psychological Society. (2002). *Code of Ethics for Psychologists Working in Aotearoa/New Zealand*. Wellington: Author.

Nuttin, J. (1955). Consciousness, behaviour, and personality. *Psychological Review*. *62*(5), 349-355.

Ones, D. S., & Andersen, N. (2002). Gender and ethnic group differences on overt personality scales in selection: Some British data. *Journal of Occupational and Organizational Psychology, 75*, 255-276.

Ones, D. S., & Viswesvaran, C. (1998). Gender, age, and race differences on overt integrity tests: Results across four large-scale job applicant datasets. *Journal of Applied Psychology, 83*, 35-42.

Packman, T., Brown, G. S., Englert, P., Sisarich, H., & Bauer, F. (2005). Differences in personality traits across ethnic groups within New Zealand and across an international sample. *New Zealand Journal of Psychology, 34*(2), 77(79).

Paunonen, S. V., & Ashton, M. C. (1998). The structured assessment of personality across cultures. *Journal of Cross-Cultural Psychology, 29*(1), 150-170.

Paunonen, S. V., & Jackson, D. N. (2000). What is beyond the Big Five? Plenty! *Journal of Personality, 68,* 821-835.

Pervin, L. A. (1990). A brief history of modern personality theory. In LA Pervin (Ed.), *Handbook of Personality: Theory and research* (pp. 3-18). New York: Guilford

Pervin, L. A. (1994). A critical analysis of current trait theory. *Psychological Inquiry, 5*, 103-113.

Pervin, L. A. (2003). *The Science of Personality* New York: Oxford University Press.

Pervin, L. A., Cervone, D., John, O. P. (2005). *Personality: Theory and Research; Ninth Edition.* New York: Wiley.

Peevers, B. H., & Secord, P. F. (1973). Developmental changes in attribution of descriptive concepts to persons. *Journal of Personality and Social Psychology, 27*(1), 120-128.

Pinker, S. (2002). *The Blank Slate: The Modern Denial of Human Nature*, Penguin Putnam.

Porter, M. E. (1985) "*Competitive Advantage*". The Free Press, New York.

Price, J. L., Monson, C. M., Callahan, K., & Rodriguez, B. F. (2006). The role of emotional functioning in military-related PTSD and its treatment. *Journal of Anxiety Disorders*, *20,* 661-674.

Raine, A., Venables, P. H., & Mednick, S. A. (1997). Low resting heart rate at age 3 years predisposes to aggression at age 11 years: Evidence from the Mauritius Child Health Project. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 1457-1464.

Rorer, L. G. (1990). Personality Assessment: A Conceptual Survey. In L. A. Pervin (Ed.), *Handbook of Personality: Theory and Research* (pp.693-720). New York: Guilford.

Roth, P.L., Van Iddekinge C.H., Huffcutt, A.I., Eidson, C.E., Jr., & Bobko, P. (2002). Correcting for range restriction in structured interview ethnic group differences: The values may be larger than we thought. *Journal of Applied Psychology*, *87,* 369-376.

Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology, 54,* 297-330.

Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human Resource Management Review, 16*(2), 155-180.

Russell, J. A., Lewicka, M., & Niit, T. (1989). A Cross-Cultural Study of a Circumplex Model of Affect. *Journal of Personality & Social Psychology, 57*(5), 848-856.

Ryan, A. M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology, 52*(2), 359-391.

Urban, S. J. (1972). *A study of the relationship of certain personality and situational variables to job satisfaction in regular elementary teachers and elementary level teachers of educable mentally retarded in second class school districts in Michigan.* Unpublished doctoral dissertation, Michigan State University.

Sackett, P. R., & Ellingson, J. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology, 50,* 707-721.

Sackett, P. R., Gruys, M. L., & Ellingson, J. E. (1998). Ability-personality interactions when predicting job performance. *Journal of Applied Psychology*. *83,* 545-556.

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative action world. *American Psychologist*, *56,* 302-318.

Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology, 76*, 323-346.

Salgado, J. F., & Fruyt, F. D. (2005). Personality in personnel selection. In A. Evers, Anderson, N., & Voskuiijl, O. (Ed.), *The Blackwell handbook of personnel selection*. Malden, MA: Blackwell Publishing.

Sampath, S. (2001). *Sampling theory and methods*. Chennai: Alpha Science International.

Saucier, G., & Goldberg, L. R. (1996). The language of personality: Lexical perspectives on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 21-50). New York: Guilford.

Schmitt, D.P., Allik, J., McCrae, R.R., Benet-Martínez, V. (2007). The geographic distribution of big five personality traits: pattern and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology*, *38*, 173-212.

Schmidt, F.L. & Hunter, J.E. (1998). The validity and utility of selection methods in personnel research: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262-274.

Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a a women? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology*. *94(1),* 168-182.

Schwab, D. P. (1971). Issues in response distortion studies of personality inventories: A critique and replicated study. *Personnel Reports, 26*, 671-675.

Sears, R. R. (1950). Personality. Annual Review of Psychology. *1*, 105-18.

Sechrest, L. (1976). Personality. *Annual Review of Psychology.* 27, 1-28.

Smith, M., George, D. (1992), "Selection methods". In Cooper, C. L., Robertson, I. T. (Eds), *International Review of Industrial and Organizational Psychology*. 7 (pp.55-97).

Snyder, M. (1983). The influence of individuals on situations: Implications for understanding the links between personality and social behaviour. *Journal of Personality,51,* 497-516.

Spearman, C. (1904). "General intelligence" objectively determined and measured. *American Journal of Psychology*, 15, 201–293.

SPSS for Windows. (2001). Rel. 11.0.1. Chicago: SPSS Inc.

Stelmack, R. M. (1990). Biological bases of Extraversion: Psychophysiological evidence. *Journal of Personality, 58,* 293-312.

Stuart, A. (1987). The ideas of sampling. 3rd ed. New York: Oxford University Press.

Tanaka, J.W., & Taylor. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitiive Psychology, 23,* 457-482.

Taylor, P., Keelty, Y., & McDonnell, B. (2002). Evolving personnel selection practices in New Zealand organisations and recruitment firms. *New Zealand Journal of Psychology, 31*(1), 8-11.

Tellegen A. (1993). Folk concepts and psychological concepts of personality and personality disorder. *Psychological Inquiry, 4*,122-130.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, *44*(4), 703-743.

Thorndike, E. L. (1921). *Educational Psychology*. New York: Teacher's College of Columbia.

Thurstone, L. L. (1947). *Multiple-Factor Analysis*. Chicago: University of Chicago Press.

Triandis, H. C. (1997). Cross-cultural perspective on personality. In Hogan, R., Johnson, J., & Briggs, S. (1997). *Handbook of personality*. San Diego: Academic Press.

Trimble, J. E. (1987). Self-perception and perceived alienation among American Indians. *Journal of Community Psychology, 15*, 316-333.

Vickers, R. R. Jr., Hervig, L. K., & Booth, R. F. (1996). *Personality and Success among Military enlisted Personnel: An historical prospective study of U.S. Navy corpsmen*. US Naval Health Research Ctr, Human Performance Dept, San. Diego, CA.

Van Dyne, L., & Saavedra, R. (1996). A naturalistic minority influence experiment: Effects on divergent thinking, conflict, and originality in work groups. *British Journal of Social Psychology, 35*, 151-167.

Wade, *T.,* Tiggemann, M.*,* Heath, A., Abraham, S.*,* Martin, N. (1995). EPQ-R personality correlates of bulimia nervosa in an Australian twin population. *Personality and Individual Differences*, *18*(2), 283-285.

Waller, N. G., Ben-Porath, Y. 1987. Is it time. for clinical psychology to embrace the five-. factor model of personality? *American Psychologist, 42*, 887-89.

Warne, E. K. (1957). *A study to determine whether the Gordon Personal Profile will tend to measure the success of student teachers*. Unpublished doctoral dissertation, Colorado State college.

Waters, L. K. (1960). *Reliability scale independence, and flight-student normes for the Gordon Personal Profile* (Research Report, Project MR005, 13-3003, Subtask 1, Report No. 32). Pensacola, Fla.: U. S. Naval School of Aviation Medicine.

Waters*,* T. J*,* (1968). The validity of the Junior Eysenck Personality Inventory Lie scale*. Educational and Psychological Measurement 28*(4), 1197-1206.

Webb, E. (1915). *Character and Intelligence*. Cambridge, Cambridge University Press.

West, S. G. (1983). Personality and prediction: An introduction. *Journal of Personality*. *51*(3), 275-285.

Whitworth, R. H., & Perry, S. M. (1990). Comparison of Anglo-and Mexican Americans on the 16PF administered in Spanish or English. *Journal of Clinical Psychology, 46*(6), 857-863.

Wiggins, J. S. (Ed., 1996). *The Five-Factor Model of Personality: Theoretical perspectives.* New York: Guilford.

Zuckerman, M. (1991). *Psychobiology of Personality.* New York: Cambridge University Press.