A study to investigate the reliability and validity of the Ten-Item Personality Inventory, when compared with two robust inventories, within a British sample

Marie Holmes

Supervised by: Jacqui Akhurst

April 2010

A study to investigate the reliability and validity of the Ten-Item Personality Inventory, when compared with two robust inventories, within a British sample

ABSTRACT

Personality inventories are commonly used instruments for the assessment of personality domains at various levels, whether it is the higher order factors or the facets underlying them. There are numerous inventories available for researchers to choose from, but a common disadvantage is the length of the inventory and consequently the completion time. The Ten-Item Personality Inventory (TIPI) aimed to address these issues. Short inventories are convenient and advantageous methodologically, but there are questions of reliability and validity, when compared with longer inventories. The aim of the current study is to investigate the validity and reliability of the TIPI, within a British sample (N = 81), when compared with the sixty item NEO-FFI and the forty-eight item EPQ-R short measure. The reliability indices of the TIPI closely corresponded to those found in the original literature. The convergent correlations indicated that the TIPI is valid within a British sample, a mean correlation of r_s = .61, compared with Furnham's (2008) r = .53. It is not recommended that the TIPI, or similar, supersede longer personality inventories, but the convenience of a short reliable and valid instrument is recognised for utilisation where time is limited or when personality is not the prime focus of research. Further psychometric investigation of the TIPI within various populations is warranted.

KEY WORDS:	PERSONALITY	TIPI	FIVE-FACTOR MODEL	RELIABILITY AND VALIDITY	BRITISH

Acknowledgements:

The dissertation process is a challenge at the culmination of any undergraduate degree programme. The ability to rise to this challenge and persevere through the inevitable highs and lows entailed, is not merely the author's quest alone. The support and guidance received from my dissertation supervisor, Dr. Jacqueline Akhurst, has been invaluable and gratefully received. The ongoing support throughout the statistical analysis process from Nathalie Noret was commendable. I would also like to recognise the support and guidance from all the lectures, especially my personal tutor, Dr. Stephen Gibson, and the technical and faculty support over the three year period that has culminated in the ability to produce a research project. Gareth Jenkins was patient and thorough in his explanation of the use of the computer programme to create my online questionnaire. Annette Webb provided support for infuriating Microsoft Word technicalities. I would also like to thank my participants for taking the time to complete the inventories. It just remains for me to thank my family, friends and peers for their unwavering support and positivity throughout the degree and in particular the dissertation process.

Introduction

Personality traits are commonly used in everyday language as descriptors of individuals, and at the other extreme, are subject to rigorous scientific investigation and empirical research. The familiarity of lexical terms utilised in daily conversation belies the complexity of academic investment in the origin and function of these basic dimensions of personality, over the past century. Personality traits are defined as 'the individual differences between people in characteristic thoughts, feelings and behaviors' (McCrae & Costa, 1995, p. 231), 'the consistencies in thought, feeling, and behavior associated with social interaction and the socioemotional aspects of life' (McAdams, 2009, p. 109) and 'the basic qualities of the person that express themselves in many contexts' (Mischel *et al*., 2003, p. 44). Many other definitions have been proffered, but the general consensus is that personality traits are dispositions that are relatively stable over time and across situations.

There are four branches of academic thinking regarding the origin and concept of traits (McAdams, 2009): the biological basis (Allport, 1937; Eysenck, 1967), the behavioural disposition (Cattell, 1957; McCrae & Costa, 1990), descriptive summaries of behavioural acts (Buss & Craik, 1983), and social constructs (Harre & Gillett, 1994; Mischel, 1996)). The theories of Eysenck (1967) and McCrae and Costa (1990) will form the basis for comparison in the current study.

The conception of traits may be as old as human language itself. The historical linear progression of personality traits can be traced from Hippocrates, to Galen, to William of Conches, to Immanuel Kant and finally to Wilhelm Wundt, the father of modern psychology (Matthews *et al*., 2003). Stelmack and Stalikas (1991) discussed the personality trait theory of Galen (A.D. 130-200), a Greek physician, based upon the theory of humours (bodily fluids). Individual differences were explained according to the combination of four humours: chole (yellow bile), melanchole (black bile), sanguis (blood) and flegma (phlegm). An individual's physical constitution and psychological characteristics were determined by their blend of humours. The legacy of Galen's theory of humours was the descriptive typology which emerged in the 18th and 19th centuries. Eysenck and Eysenck (1985) state that temperament traits can be traced back to Galen's four temperament system of humours.

The contemporary approach to trait personality theory originated with Allport (1937). He adopted an idiographic methodology, examining case-studies and analysing interviews; he was interested in describing personality, rather than probing personality. Allport proposed a distinction between traits, he categorised them into cardinal, central and secondary traits. His idiographic approach highlighted the uniqueness of personality, but his work with cardinal traits influenced the work of later theorists who adopted a nomothetic approach. The utilisation of statistical techniques and rigorous scientific procedures in the measurement of personality traits was initiated by Cattell (1957), with the analysis of 4504 words originally identified by Allport & Odbert (1936), reducing them to 16-personality factors.

The scientific study of traits required systematic data collection, statistical techniques for data analysis and the development of testable theories. Theorists constructed hypotheses relating to the number and nature of personality traits, and designed questionnaires to measure them (Larsen & Buss, 2005). Psychometric principles were adopted to investigate the effectiveness of the measuring tool, and to address

any modifications considered necessary. Any given single trait measure must meet three essential criteria: reliability, stability and validity (Coolican, 2006). Reliability refers to the accuracy with which a questionnaire measures the quality it is pertaining to measure. The consistency can be measured internally or externally. Internal reliability is the assessment of the consistency of a scale within itself, do the items within a questionnaire measure the same thing. Cronbach's (1951) alpha is used to measure the internal reliability of a questionnaire; the alpha tends to increase as the inter-item correlation increases and as the number of items within a tool increases. The internal reliability of the three psychometric tools forming the focus of the current study will be analysed, and compared with existing research. External reliability refers to the stability of a psychometric tool, and is measured utilising test-retest reliability. The validity of a personality questionnaire, questions that the tool is actually measuring what it purports to measure. Construct validity is the ultimate goal for theory driven research. It encompasses all of the empirical evidence and theoretical analysis surrounding a trait, referred to as the 'nomological network' (Eysenck, 1957, p.261). Construct validity is in flux as new research evidence is considered, resulting in the modification of hypotheses and concepts.

Reliability, stability and validity are useful when measuring single traits, for example Extraversion or Conscientiousness, but when considering multiple traits as the model of personality they are not sufficient. Factor analysis (Guilford & Guilford, 1934) is the method utilised to simultaneously identify multiple traits by inputting of all the correlations between all of the items comprising a questionnaire. The responses are homogenously clustered, simplifying the correlation matrix and identifying underlying factors which account for the variation in individuals' scores. The resulting unrelated dimensions are identified as higher order traits, know as source traits - sixteen (Cattell, 1957), supertraits - three (Eysenck, 1967) or the Big Five (Costa & McCrae, 1985). It is the responsibility of the researcher to identify the number of and allocate names to the different traits, disagreement has evolved and an example will be discussed at greater length, focussing on the debate between Eysenck's (1992) Psychoticism and Costa and McCrae's (1992b) Agreeableness and Conscientiousness.

Debate surrounds the utilisation of the statistical procedure of factor analysis for identification of the number of personality traits (Block, 2001); although the utility of reducing a large data set to one more manageable is recognised. There is a degree of subjectivity involved in choosing the items to be analysed, if components of a trait are not entered for analysis, that trait will not form a resultant factor (Larsen & Buss, 2005). The factors to emerge are purely descriptive and provide no theoretical support or advancement (Block, 2001). All three inventories utilised in the current study have their origins in factor analysis.

Eysenck (1967) was highly influential in the trait approach to personality and was concerned with developing quantitative methods that enabled observable personality variations to be assessed using robust statistical procedures. He developed and adapted several personality questionnaires over a forty year period. His trait personality theory was heavily grounded in psychophysiology, and he believed traits to be highly heritable. He named his three supertraits: Extraversion, Neuroticism and Psychoticism. The biological basis of the traits varies, Extraversion is considered to be linked to the cortical arousal system in the brain, and Neuroticism is considered to be linked to the activation system in the sympathetic nervous system (Eysenck,

1967). The empirical support for Psychoticism is less robust, but there is some indication that it may be linked to testosterone levels (Furnham *et al*., 2008). Eysenck (1967) developed and modified several questionnaires, primarily utilising factor analysis, in his efforts to create a reliable and valid trait assessment tool. The Eysenck Personality Inventory (EPI: Eysenck & Eysenck, 1964), a fifty-seven item self-report measure, which included a Lie scale, developed out of the Maudsley Medical Questionnaire (Eysenck, 1947). The EPI was written in more simplified terms to make it more accessible to the general population. There was no correlation between Extraversion and Neuroticism, which had been found within an earlier questionnaire. The Eysenck Personality Questionnaire (EPQ: Eysenck & Eysenck, 1975), a ninety item self-report measure is considered to be the most important and frequently used Eysenkian measure (Furnham *et al*., 2008). The EPQ also contained a new scale to measure Psychoticism, and there was a shift in terms from Neuroticism and Psychoticism, to Emotionality and Tough-Mindedness, a conscious effort by Eysenck to move away from psychiatric terms.

The Eysenck Personality Questionnaire Revised (EPQ-R: Eysenck *et al*., 1985), a one hundred item self-report measure, attempted to address the issues of the Psychoticism scale. The Extraversion and Neuroticism scales correlate highly across all measures, and are generally found to be normally distributed among the population, with acceptable coefficient alphas (above 0.7) (Kline, 2000). However, the Psychoticism scale has its problems, the distribution tends to be skewed towards low scores, and in males, there tends to be a slight positive correlation between the Psychoticism scale and the other two scales, low coefficient alphas tend also to result (e.g. Alexopoulos & Kalaitzidis, 2004; Aluja *et al*., 2003; Francis *et al*., 2006). Reluctance of individuals who may score highly on the Psychoticism scale to participate in research is cited as a possible explanation for the low scores often found (Eysenck *et al*., 1985). Ongoing modification of the Psychoticism scale has been recommended (Furnham *et al*., 2008). A short form (EPQ-R short measure: Eysenck *et al*., 1985), a forty-eight item self-report measure, twelve items per scale, was devised for when time is limited; it is this questionnaire that will be utilised in the current study. The coefficient alphas for the EPQ-R short measure reported by Eysenck *et al*. (1985) for males and females respectively are 0.76 and 0.78 for Psychoticism, 0.85 and 0.90 for Extraversion, 0.85 and 0.88 for Neuroticism and 0.79 and 0.82 for the Lie scale.

Eysenck's (1967) three supertraits have been challenged by the Big Five, the most empirically endorsed being Extraversion, Neuroticism, Agreeableness, Conscientiousness and Openness (Costa & McCrae, 1985). Digman (1990) claims that over forty years of research literature points towards five personality domains, with debate surrounding the names of the domains. Costa and McCrae (1985) utilised the domains explained by Norman (1963), from his analysis of peer ratings; Extraversion (Surgency) was the same, Emotional Stability equated with Neuroticism, and Culture equated with Openness. However, their initial inventory did not contain Agreeableness and Conscientiousness; therefore they constructed brief scales to measure them, and published the NEO-Personality Inventory (NEO-PI: Costa & McCrae, 1985) of one-hundred and eighty-one items. Further research was undertaken to develop facet scales for the Agreeableness and Conscientiousness domains, resulting in the NEO-Personality Inventory-Revised (NEO-PI-R: Costa & McCrae, 1992a), with two-hundred and forty items. The NEO-PI and the NEO-PI-R examine the five factors at facet level; each factor is measured on six sub-scales.

It was considered that a short measure would provide a global view of an individual; Costa and McCrae (1992a) developed the NEO-Five Factor Inventory (NEO-FFI) to provide this view. The NEO-FFI comprises sixty items, twelve relating to each factor, identified via factor analysis of an earlier administration of the NEO-PI. The coefficient alphas reported by Costa and McCrae are 0.86 for Neuroticism, 0.77 for Extraversion, 0.73 for Openness, 0.68 for Agreeableness and 0.81 for Conscientiousness. The NEO-FFI will form the second inventory for comparison in the current study. Costa and McCrae (1985) have developed robust psychometric tools which demonstrate validity and replicability, however, there is little empirical research explaining the basis of the five factors, although McCrae and Costa (1990) do suggest a biological basis (heritability). This is debated and some view the five-factor approach purely as a descriptive framework (e.g. Block, 1995, 2001; McAdams, 1992).

Eysenck's (1991, 1992) Psychoticism scale has been compared to the combination of Costa and McCrae's (1992b) Agreeableness and Conscientiousness (reversed) scales. Eysenck (1992) argues that the taxonomy of a field of study needs to be presented before the causal theories can be established; adding that the two are mutually beneficial and refinement may occur as a result of empirical study. Costa and McCrae (1992b) proposed five rationales for the existence of five factors of personality. Firstly, longitudinal (e.g. Costa & McCrae, 1988) and cross-sectional studies (e.g. Allemand *et al*., 2008) had shown five robust factors to be enduring behavioural dispositions. Secondly, traits associated with the five factors had emerged from different personality systems and from lexical studies. Thirdly, the five factors have been demonstrated across different ages (e.g. Allemand *et al*., 2008; Terracciano *et al*., 2006), sexes (e.g. Feingold, 1994), races and languages (e.g. McCrae *et al*., 1999). Fourthly, a biological basis for each of the five factors has been demonstrated thorough heritability studies. Finally, cross-cultural studies (e.g. Costa *et al*., 2001) have shown similarities in the ageing trajectories of the five factors, suggesting they are universal. Eysenck (1992) agreed that the criteria set out by Costa and McCrae (1992b) were necessary for accepting the five factor model, but insufficient for determining the important dimensions of personality. Eysenck (1992) argued that Agreeableness and Conscientiousness were in fact facets of his higher-order factor Psychoticism. The relationship of Psychoticism (EPQ-R short measure) with Agreeableness and Conscientiousness (NEO-FFI and TIPI) will be examined in the current study. Openness has also been linked to intellect (Peabody & Goldberg, 1989). Eysenck (1967) was strict in establishing the nomological network of a trait dimension, and believed it was this that provided psychometric validity. This network he felt was missing in the five-factor model and consequently he questioned its psychometric validity.

One of the prime assumptions of a personality trait is that it is stable over time. The big five have been rigorously investigated for their stability across time utilising longitudinal studies (e.g. Terracciano *et al*., 2006). According to McCrae *et al*. (1999) consensus now appears to be that between the ages of eighteen and thirty, the mean trait levels of Neuroticism, Extraversion and Openness decrease slightly, whereas Agreeableness and Conscientiousness increase slightly. After the age of thirty mean personality traits remain relatively stable. More recently Allemand *et al*. (2008) reported findings that mean levels of Agreeableness and Conscientiousness continue to rise into old age, with their oldest participants demonstrating higher mean levels than the participants in the young and middle age groups. The current study

will examine the pattern of mean trait levels across age groups and identify whether it is consistent across all three inventories.

It was considered that personality was formed by the onset of adulthood and remained constant, but Erikson (1966) disputed this assumption. Erikson's theory was utilised to determine the three age groups considered in the current study. The sixth stage denotes intimacy versus isolation, young adulthood to early middle age (18-35 years), when individuals are building relationships and creating a sense of belonging outside of the family circle, Erikson felt that failure to establish these relationships may lead to feelings of isolation. The seventh stage denotes generativity versus self-absorption, middle age (36-59 years), signifying a shift from focussing on the immediate family circle to an increased concern regarding the population at large and a desire to help improve circumstances for others, especially future generations. A failure to recognise the needs of individuals and society beyond one's own dominant concerns Erikson termed self-absorption or stagnation. The eighth stage denotes integrity versus despair, late middle age onwards (60-100 years), signifying a change of pace. Individuals who reflect and consider their lives to have been productive and worthwhile experience a sense of integrity, whereas individuals who reflect and feel disappointment and failure to have permeated their lives, may feel time is limited to readdress the balance, and experience despair. Demonstration of certain personality traits have been linked to the dichotomies evident in Erikson's stages (Matthews *et al*., 2003).

Sex differences in personality traits may be expected and explained in terms of biological/ evolutionary or social models of personality (Matthews *et al*., 2003). Two meta-analyses (Feingold, 1994; Costa *et al*., 2001) have indicated that there are differences in personality traits between males and females, predominantly in Extraversion, Neuroticism and Agreeableness. Males were generally found to score higher on Extraversion (depending on the inventory completed) and females were generally found to score higher on Neuroticism and Agreeableness. It is surmised that gender differences in personality traits tend to follow gender stereotypes, but the differences found are small compared with intra-group differences (Matthews *et al*., 2003).

Attempts have been made to produce a short personality inventory, measuring global personality traits; for example the EPQ-R short measure (Eysenck *et al*., 1985) with forty-eight items has a completion time of 15 minutes. However, it was felt that a very brief measure was required, especially for use in clinical settings where time can be limited, and for when personality was not the foremost area of interest. The Ten-Item Personality Inventory (TIPI: Gosling *et al*., 2003) was developed to fulfil this role. The authors based their research on the five-factor approach to personality and compared a five item measure and a ten item measure, with the Big-Five Inventory (BFI: Benet-Martínez & John, 1998), consisting of forty-four items. The new ten item measure was constructed using two descriptors representing the five domains, one item for each extreme. Convergent and discriminant validity, and reliability, via test-retest analysis, were examined. The internal reliabilities, measured using Cronbach's alpha, were relatively low: 0.68 for Extraversion, 0.40 for Agreeableness, 0.50 for Conscientiousness, 0.73 for Emotional Stability (Neuroticism reversed) and 0.45 for Openness to Experience. The negative aspect of a brief inventory is considered to be its psychometric properties, and the authors recognise that the reliability alphas will be low and the confirmatory and exploratory

factor analyses indices will be poor. The internal reliability of the TIPI will be assessed in the current study and compared with the alphas found in the original study; a principal component analysis will also be performed.

The participants in the Gosling *et al*. (2003) study were 1813 undergraduate psychology students, and a sub-sample of 180 completed the inventories again six weeks later, providing the test-retest data. The convergent correlations were substantial (0.65 - 0.87) and the discriminant correlations did not exceed 0.36. The authors state that the TIPI takes one minute to complete and provides 'a reasonable proxy for longer Big-Five instruments' (p.523). The authors also favour the increased use of the TIPI in future research as a method of accumulating knowledge about its psychometric properties. Several researchers (e.g. Ehrhart, 2009; Herzberg, 2006; Hofmans, 2008; Muck, 2007), in Europe and the USA, have investigated the psychometric properties of the TIPI and the general consensus appears to be that the TIPI is a valid instrument for assessing broad personality domains when available time is short.

Various studies have compared different inventories, many based in the USA (e.g. Zuckerman *et al*., 1993) or mainland Europe (e.g. Barelds & Luteijn, 2002), but from a literature search none have been identified comparing the three inventories in the current study. Furnham (2008) investigated the intercorrelations between the NEO-FFI, the TIPI, a five-item Single-Item Measure of Personality (SIMP: Woods & Hampson, 2005) and self-estimated personality (Furnham & Chamorro-Premuzic, 2004). He notes the variations in length of inventories and attributes these to measures of traits at the domain level or the facet level, the desire for high internal reliability hence validity, and whether the inventory is ipsative or nonipsative. The study involved one hundred undergraduate participants, predominantly female (78%) and of white British origin (84%). The results demonstrated all thirty correlations as significant and positive. The correlations between the NEO-FFI and the TIPI were 0.61 for Neuroticism, 0.48 for Extraversion, 0.52 for Openness, 0.39 for Agreeableness and 0.66 for Conscientiousness. Furnham (2008) concludes that the TIPI demonstrates 'slightly better' (p. 315) validity than the other two measures, and argues that if the predictive validity of a short inventory is equal to that of a longer inventory, justification for its use can be made.

Herzberg and Brähler (2006) compared the NEO-FFI and TIPI in a German sample. They also included a revised short form containing sixteen adjectives, measuring the big five. The study involved a first sample of 2552 (53% female) participants aged 14-99 years representing the general German population, and a second sample of 200 females and 164 males aged 18-94 years, who were friends or relatives of undergraduate students. The internal reliabilities of the TIPI were 0.54 for Neuroticism, 0.24 for Extraversion, 0.41 for Openness, 0.33 for Agreeableness and 0.52 for Conscientiousness. The convergent correlations between the NEO-FFI and the TIPI were 0.66 for Neuroticism, 0.45 for Extraversion, 0.23 for Openness, 0.08 for Agreeableness and 0.46 for Conscientiousness. The authors reported improved reliability and convergent correlations for their revised short form, when compared to the TIPI, and subsequently recommend the use of their brief adjective measure in preference to the TIPI, when the use of a longer Big-Five measure is not appropriate.

The research questions examined in the current study will focus on:

- The psychometric properties of three inventories, specifically whether the TIPI is a valid and reliable measure of personality traits, when compared with two established inventories.
- The internal reliability of the three measures within the population sampled, and the comparison to published data.
- The relationship between inventories, examining the correlations, whether items are measuring the trait they purport to be measuring; and examination of any variation in the sample by sex and age group.
- To identify whether the TIPI is a valid and reliable measure of the Big-Five personality traits, within a British sample. A British sample has been chosen since minimal research has been identified utilising the TIPI within a British sample.
- A comparison between Eysenck's Psychoticism scale (EPQ-R short measure) and Costa and McCrae's Agreeableness and Conscientiousness scales (NEO-FFI and TIPI) to examine the debate that Psychoticism is an amalgamation of Agreeableness and Conscientiousness.

Method:

Design:

This is an exploratory study, therefore a cross-sectional design was chosen, involving correlational analyses. Participants were assigned to groups according to demographic variables, for aspects of the study.

Participants:

In an attempt to recruit a heterogeneous sample, a snowball recruitment method was utilised. Eighty-five contacts, meeting the research criteria (adult (18+) and British currently residing in the UK), were initially emailed and requested to complete the inventories and forward them on to family and friends also meeting the research criteria. No incentives were offered for participation, excepting a pledge to deliver a summary of the research findings on completion of the study. The necessity of British citizenship was highlighted due to the focus being on a British sample.

The participants comprised 81 British citizens, currently residing in the United Kingdom. Specifically, females 74.1 % (n = 60) and males 25.9 % (n = 21); White 96.3 % (n = 78), Mixed Race 2.5 % (n = 2) and Asian British 1.2 % (n = 1). No participants reported Black or Chinese ethnicity, and all divulged their ethnic status, sex and age. Average age of the whole sample was 32.98 years (SD = 15.36 years), males 34.52 years (SD = 16.35 years) and females 32.43 years (SD = 15.10 years).

Measures:

The psychometric instruments used in this study were all self-report inventories: the Ten-Item Personality Inventory (TIPI: Gosling *et al*., 2003), the Eysenck Personality Questionnaire-Revised short measure (EPQ-R short measure: Eysenck *et al*., 1985) and the NEO-Five Factor Inventory (NEO-FFI: Costa & McCrae, 1992a). Other instruments were considered but rejected for various reasons.

Ten-Item Personality Inventory (TIPI):

The TIPI (Gosling *et al*., 2003) contains ten statements, two of the statements each relate to one of the five dimensions of the five factor model (Extraversion, Agreeableness, Conscientiousness, Emotional Stability and Openness). Emotional Stability is measured on this scale so that higher scores indicate higher Emotional Stability, whereas in measures of Neuroticism higher scores indicate lower Emotional Stability. A high score on the TIPI Emotional Stability scale would possibly be indicative of a low score on the NEO-FFI and EPQ-R short measure Neuroticism scales.

Each item is rated on a 7-point Likert scale: 1 = Disagree strongly, 2 = Disagree moderately, 3 = Disagree a little, 4 = Neither agree nor disagree, 5 = Agree a little, 6 = Agree moderately and 7 = Agree Strongly. The items are 'extraverted, enthusiastic' and 'reserved, quiet' (reverse-scored) for Extraversion, 'critical, quarrelsome' (reverse-scored) and 'sympathetic, warm' for Agreeableness, 'dependable, self-disciplined' and 'disorganised, careless' (reverse-scored) for Conscientiousness, 'anxious, easily upset' (reverse-scored) and 'calm, emotionally stable' for Emotional Stability, and 'open to new experiences, complex' and 'conventional, uncreative' (reverse-scored) for Openness. The score on each dimension denotes the extent to which the individual reports that particular trait.

Eysenck's Personality Questionnaire-Revised (EPQ-R) short measure:

The EPQ-R short measure (Eysenck *et al*., 1985) contains forty-eight items requiring a YES or NO response. The items are split into four sets of twelve covering three personality dimensions: Extraversion, Neuroticism and Psychoticism. The fourth Lie scale will not be considered in the current study.

Examples of the items in the EPQ-R short measure include:

Extraversion:	3. Are you a talkative person?
	7. Are you rather lively?
Neuroticism:	1. Does your mood often go up and down?
	5. Do you ever feel 'just miserable' for no reason?
Psychoticism:	2. Do you take notice of what people think?
	6. Would being in debt ever worry you?
Lie:	4. If you say you will do something, do you always keep your promise no matter how inconvenient it might be?
	8. Were you ever greedy by helping yourself to more than your fair share of anything?

The questionnaire is scored by allocating one point to each answer, correlating with a key. On some questions (e.g. items 1, 7 and 22) a YES response gains a point, but on other questions (e.g. items 6, 27 and 43) a NO response gains a point. Each scale is scored out of twelve; higher scores indicate the degree to which an individual reports that particular personality trait.

NEO – Five Factor Inventory (NEO-FFI):

The NEO-FFI (Costa & McCrae, 1992a) contains five, twelve item scales measuring the Big Five domains of personality: Neuroticism, Extraversion, Openness, Agreeableness and Conscientiousness. Each item is rated on a 5-point Likert scale: 0 = Strongly Agree, 1 = Agree, 2 = Neutral, 3 = Disagree and 4 = Strongly Disagree. Thirty-three of the items are reverse-scored.

Examples of items in the NEO-FFI include:

Neuroticism:	1. I am not a worrier.
	6. I often feel inferior to others. (reverse-scored)
Extraversion:	2. I like to have a lot of people around me. (reverse-scored)
	7. I laugh easily. (reverse-scored)
Openness:	3. I don't like to waste my time daydreaming.
	8. One I find the right way to do something, I stick to it.
Agreeableness:	4. I try to be courteous to everyone I meet. (reverse-scored)
	9. I often get into arguments with my family and co-workers.
Conscientiousness:	5. I keep my belongings clean and neat. (reverse-scored)
	10. I'm pretty good at pacing myself so as to get things done on time. (reverse-scored)

Higher scores denote the extent to which an individual demonstrates that particular personality trait.

The computer programme utilised to create the online questionnaire was Survey software SelectSurvey.NET 2.3.11. The individual scores were available in the format of a Microsoft Excel spreadsheet. All statistical analysis was computed utilising SPSS version 17.0.

Procedure:

The purpose of this study was to compare a short measure of personality traits with longer more established inventories, to examine its reliability and validity, within a British sample. The choice of short inventory (i.e. TIPI) was relatively straightforward, as only one brief study (i.e. Furnham, 2008) was found examining the relationship between the TIPI and other inventories, within a British student population. Several comparative inventories were considered, including the Sixteen Personality Factor Questionnaire (16-PF: Cattell *et al*., 1993), the Myers-Briggs Type Inventory (MBTI: Myers *et al*., 1998), the Minnesota Multiphasic Personality Inventory (MMPI: Hathaway & McKinley, 1967) and the Keirsey Temperament Sorter (Keirsey, 1998). These were rejected for reasons of length, complexity of scoring and the requirement for specialised training. The two comparison inventories chosen were the NEO-FFI (Costa & McCrae, 1992a) and the EPQ-R short measure (Eysenck *et al*., 1985). The NEO-FFI provided a direct comparison with the same personality traits, and the EPQ-R short measure focused on two comparable scales (Extraversion and Neuroticism/ Emotional Stability) and also measured on the Psychoticism scale,

which has been compared to the Agreeableness and Conscientiousness scales (Costa & McCrae, 1992b; Eysenck, 1992).

The proposal was presented to the Ethics Committee for consideration, and received ethical approval following further clarification on the mode of recruitment. The three inventories were presented online, utilising the Survey software SelectSurvey.NET 2.3.11 package. Eighty-five individuals were emailed requesting their participation and giving a link to the online survey. All individuals were requested to forward the email on to friends and family also meeting the research criteria, to increase the range of individuals sampled. An issue to note was that it is not possible to be certain all participants are eighteen plus and resident British citizens, reliance on the honesty of the participant is necessary.

The informed consent form was the initial page presented once a participant clicked on the link to open the survey. Participants were assured regarding confidentiality and anonymity issues, and data protection. They were informed that they could withdraw from the study at any time, and how to achieve this. It was also made clear that compulsory completion of all questions on all three inventories was not essential. Contact information of various organisations (e.g. Samaritans, NHS Direct) was also provided, in the eventuality of stress or anxiety being provoked by items during the completion of the inventories. Contact details of the researcher were also provided for clarification of any questions.

The second page of the online survey requested background information. Participants were requested to provide a password, for use if they wished to withdraw from the study. Information regarding sex, age and ethnicity was also collected. If the participant decided to proceed and complete the inventories, the next page revealed the TIPI, followed by the EPQ-R short measure on the fourth page and finally the NEO-FFI on the final page. A message thanking the participant for completing the survey was posted on completion. Presentation order of the three inventories was stable across all participants; it was considered that it may be preferable to vary the order of presentation between participants to allow for fatigue effects, however, the survey software package would not allow this. The final day for online survey completion was the last day of January 2010. All data obtained were saved to a Microsoft Excel file and entered into an SPSS data file for analysis.

All data were entered into SPSS and missing data were addressed. Any missing data on the NEO-FFI and the EPQ-R short measure were amended by utilising the mean of the remaining scores for that particular individual on the corresponding scale. All items requiring reverse-scoring on the TIPI and the NEO-FFI were then addressed, and the items on all three inventories were sub-scaled producing the individual scores for each trait, for each participant. Data analysis commenced with the running of descriptive statistics in percentages per category, and the mean (SD) scores for the whole sample, sex and age were reported. Tests of normality were run and analysed, assessing whether the data was parametric or non-parametric.

Cronbach's alpha statistic was utilised to measure the reliability indices of the three inventories. A principal component analysis with varimax rotation was undertaken on the TIPI scores, in order to examine the factor loadings. Correlations between all trait scores were analysed using Spearman's Rho, due to the data being non-parametric. The data file was split by sex and age group and the correlations computed.

Correlations between the EPQ-R short measure Psychoticism scores and the TIPI and NEO-FFI Agreeableness and Conscientiousness scores were examined using Spearman's Rho, due to the data being non-parametric. The data file was split be sex, age, and sex and age in order to examine the various correlations at sub-group level.

Debriefing the participants involved the publishing of contact information, should stress or anxiety have been evoked, and provision of researcher contact details for clarification of any questions. In addition, upon completion of the study a summary of the research findings will be emailed to all eighty-five original contacts and they will be requested to forward it on to the individuals they had previously contacted to participate in the study.

Results:

Descriptive analysis:

Descriptive statistics, specifically mean and standard deviation were computed for all scores across all three inventories. The sample as a whole was computed (see Table 1), followed by splitting the data file for sex (see Table 1) and age (see Table 2), and finally descriptive statistics for the split of sex and age were computed.

Table 1

Descriptive statistics for the total sample and split by sex across all measures. ***Mean (SD)***

	No. Items	Total (N=81)	Males (n=21)	Females (n=60)
TIPI				
Emotional Stability	2	9.05 (3.06)	10.43 (2.75)	8.57 (3.04)
Extraversion	2	9.11 (3.23)	9.81 (2.98)	8.87 (3.30)
Openness	2	10.11 (2.41)	10.67 (2.54)	9.92 (2.36)
Agreeableness	2	10.12 (2.09)	9.81 (2.04)	10.23 (2.11)
Conscientiousness	2	10.44 (2.57)	11.00 (2.53)	10.25 (2.58)
NEO-FFI				
Neuroticism	12	24.42 (10.14)	19.67 (9.71)	26.08 (9.83)
Extraversion	12	28.78 (6.37)	30.81 (6.66)	28.07 (6.16)
Openness	12	27.58 (6.06)	24.86 (6.64)	28.53 (5.59)
Agreeableness	12	31.53 (6.67)	29.95 (6.49)	32.08 (6.70)
Conscientiousness	12	30.98 (7.22)	31.86 (5.70)	30.67 (7.70)
EPQ-R short measure				
Neuroticism	12	6.31 (3.72)	4.86 (4.22)	6.82 (3.42)
Extraversion	12	7.69 (3.65)	8.24 (3.42)	7.50 (3.73)
Psychoticism	12	1.86 (1.47)	2.33 (1.59)	1.70 (1.41)

TIPI – Ten Item Personality Inventory NEO-FFI – NEO Five Factor Inventory

EPQ-R short measure – Eysenck's Personality Questionnaire Revised – short measure

The whole sample mean scores for the TIPI were comparable with those found in the Furnham (2008) study, which ranged from 10.4 for Openness to 8.6 for Neuroticism, with the rank order of the mean scores corresponding, with the exception of the reversal of the Openness and Conscientiousness positions. The mean scores for

the NEO-FFI were lower than those found in the Furnham (2008) study. It was not possible to directly compare the mean scores across inventories, due to the inventories measuring on different scales; therefore mean scores were ranked for each inventory, from highest to lowest, and compared. Agreeableness and Conscientiousness ranked highest for the TIPI and NEO-FFI and Psychoticism ranked lowest in the EPQ-R short measure. This will be discussed with reference to the Psychoticism (Eysenck, 1992) or Agreeableness and Conscientiousness (Costa & McCrae, 1992b) debate. No other consistent ranking pattern was found in the sample as a whole.

A similar ranking pattern was identified for the females regarding the Psychoticism and Agreeableness/ Conscientiousness mean scores, but the pattern was not repeated in the male sample (see Table 1). No other consistent ranking pattern was identified across the measures when the sample was split by sex.

Table 2

Descriptive statistics for the three age groups across all measures. *Mean (SD)*

	18-35 years (n=51)	**36-59 years (n=24)**	**60-100 years (n=6)**
TIPI			
Emotional Stability	8.78 (3.12)	9.38 (2.48)	10.00 (4.69)
Extraversion	9.08 (3.33)	8.92 (3.20)	10.17 (2.64)
Openness	9.94 (2.60)	10.29 (2.03)	10.83 (2.32)
Agreeableness	10.00 (1.99)	10.54 (2.38)	9.50 (1.64)
Conscientiousness	9.98 (2.52)	10.96 (2.58)	12.33 (2.07)
NEO-FFI			
Neuroticism	25.96 (10.04)	22.54 (9.50)	18.83 (11.96)
Extraversion	28.63 (7.05)	28.96 (4.68)	29.33 (7.12)
Openness	27.80 (6.53)	27.54 (5.43)	25.83 (4.58)
Agreeableness	30.84 (6.76)	33.25 (6.44)	30.50 (6.63)
Conscientiousness	29.20 (6.49)	33.63 (7.34)	35.50 (8.64)
EPQ-R short measure			
Neuroticism	6.55 (3.51)	6.04 (3.76)	5.33 (5.57)
Extraversion	7.76 (3.94)	7.54 (3.37)	7.67 (2.25)
Psychoticism	2.14 (1.33)	1.38 (1.74)	1.50 (1.05)

TIPI – Ten Item Personality Inventory NEO-FFI – NEO Five Factor Inventory

EPQ-R short measure – Eysenck's Personality Questionnaire Revised – short measure

The ranking pattern for the Psychoticism and Agreeableness/ Conscientiousness scales continued across the three age groups, with the exception of the oldest age group, where the Agreeableness mean score ranked lowest of all five mean scores (see Table 2). The NEO-FFI demonstrated an identical ranking pattern across all age groups, with the exception of the reversal of the ranking positions for the Agreeableness and Conscientiousness mean scores in the youngest age group (see Table 2). No other consistent ranking pattern was identified across the measures when the sample was split by age group.

When the sample was sub-divided by age and sex no ranking pattern was identified. The EPQ-R short measure maintained the same ranking pattern across the whole sample and all sub-samples; it was the only inventory to do so.

Principal component analysis:

The linear component structure of the TIPI was analysed by means of a principal component analysis, with varimax rotation, of the TIPI scores. The Kaiser measure of sampling adequacy was 0.69, and the average communality was 75%, explaining the amount of variance accounted for by the analysis. The analysis was run on five components in order to reflect the five personality traits purportedly measured by the TIPI, albeit an eigenvalue >1 was only reported for three factors (2.91, 1.79 and 1.20 respectively), an eigenvalue <1 was reported for the remaining two factors (0.84 and 0.82 respectively). There is debate surrounding the level of the eigenvalue for a factor to be retained; Kaiser (1960) reports the prerequisite of a value >1, whereas Jolliffe (1972) states >.7 is sufficient for a factor to be retained. If Jolliffe's argument is to be followed all five factors should be retained.

Table 3

TIPI Principal Component Analysis with varimax rotation.

TIPI Questions	Component				
	1	2	3	4	5
1. Extraversion - extraverted, enthusiastic	.859	.093	.007	-.012	.183
2. Agreeableness – critical, quarrelsome	-.336	.688	-.157	.248	.101
3. Conscientiousness – dependable, self-disciplined	.151	.564	.351	.299	-.421
4. Emotional Stability – anxious, easily upset	.352	.405	.454	-.501	.226
5. Openness – open to new experiences, complex	.197	.173	.243	.128	.816
6. Extraversion – reserved, quiet	.856	-.081	.142	.015	.004
7. Agreeableness – sympathetic, warm	.046	.164	.056	.898	.119
8. Conscientiousness – disorganised, careless	-.090	.243	.772	-.071	.038
9. Emotional Stability – calm, emotionally stable	.158	.798	.277	-.082	.177
10. Openness – conventional, uncreative	.303	-.078	.730	.117	.193

TIPI – Ten Item Personality Inventory

Table 3 shows that component one has high positive loadings for questions 1 and 6, indicating the factor is loading on Extraversion, commensurate with the TIPI scale. The remaining four components fail to demonstrate high factor loadings, as cautioned by Gosling *et al*. (2003).

Cronbach's Alpha:

Cronbach's alpha statistic was utilised to examine the reliability indices of all three inventories. Resulting alphas were compared with the alphas reported in the literature relating to the development of the inventories (see Table 4).

Table 4

Reliability indices for all measures calculated using Cronbach's Alpha, and the comparable alphas from the classic literature.

	α **N = 81**	**A** **Gosling *et al*. (2003)** **N = 1813**	**α** **Costa and McCrae (1992a)** **N = 732**	**α** **Eysenck *et al*. (1985)** **N = 902**
TIPI				
Emotional Stability	.68	.73		
Extraversion	.73	.68		
Openness	.46	.45		
Agreeableness	.37	.40		
Conscientiousness	.36	.50		
NEO-FFI				
Neuroticism	.90		.81	
Extraversion	.77		.71	
Openness	.70		.68	
Agreeableness	.83		.65	
Conscientiousness	.86		.69	
EPQ-R short measure				
Neuroticism	.87			.87
Extraversion	.88			.88
Psychoticism	.38			.77

TIPI – Ten Item Personality Inventory NEO-FFI – NEO Five Factor Inventory

EPQ-R short measure – Eysenck's Personality Questionnaire Revised – short measure

The reliability indices of the NEO-FFI and EPQ-R short measure were high within the population sampled, with all alphas being .7 or above (Kline, 2000), with the exception of the Psychoticism scale on the EPQ-R short measure. The TIPI demonstrated moderate alphas (.68 - .73) for the Emotional Stability (Neuroticism reversed) and Extraversion scales but for the remaining three scales the alphas were low (.36 - .46). The EPQ-R short measure alphas for the Neuroticism and Extraversion scales compared exactly with the original alphas (Eysenck *et al*., 1985) but there was a large discrepancy between the Psychoticism alphas. The NEO-FFI compared closely for the Neuroticism, Extraversion and Openness scales in terms of alphas, but the Agreeableness and Conscientiousness scales produced higher alphas in the current study than those reported in the original literature (Costa & McCrae, 1992a). The TIPI followed a similar pattern to the NEO-FFI with the alphas reported for the Emotional Stability (Neuroticism reversed), Extraversion and Openness scales corresponding closely with those reported in the original paper (Gosling *et al*., 2003). The Agreeableness alpha was also similar to the one originally

reported, but the Conscientiousness scale produced an alpha markedly lower than that reported by Gosling *et al*.

Correlations:

To examine the relationship between the TIPI, the NEO-FFI and the EPQ-R short measure correlations were computed. Scatter graphs for all significant correlations were drawn and examined. Data were found to be non-parametric, therefore correlations were analysed using Spearman's Rho (see Table 5).

The expected correlations were significant and ranged from .42 to .81. A strong negative correlation was found between the TIPI Emotional Stability (Neuroticism reversed) scores and the NEO-FFI and EPQ-R short measure Neuroticism scores. The TIPI Extraversion scores correlated most strongly with the EPQ-R short measure, but there was also a strong correlation with the NEO-FFI Extraversion scores. The TIPI Openness scores correlated significantly with seven out of the ten possible scales, the highest being with the NEO-FFI Openness scores. The TIPI Agreeableness scores significantly correlated with the NEO-FFI Agreeableness scores only. The TIPI Conscientiousness scores demonstrated a strong positive correlation with the NEO-FFI Conscientiousness scores, and also significantly correlated with three other scales, two negatively and one positively. The NEO-FFI Neuroticism and Extraversion scores demonstrated strong positive significant correlations with the corresponding EPQ-R short measure scores. The correlations reported between the TIPI and NEO-FFI corresponding scales were higher than those reported in the Furnham (2008) study, with the exception of the Openness scale. The mean convergent correlation in the present study is r_s = .61, compared to r = .53 in the Furnham study. All remaining correlations were non-significant and will not be discussed further.

Table 5

Spearman's Rho Correlations, between all three inventories, across all scales. (N = 81)

	TIPI-ES	TIPI-E	TIPI-O	TIPI-A	TIPI-C	EPQ-N	EPQ-E	EPQ-P	NEO-N	NEO-E	NEO-O	NEO-A	NEO-C
TIPI-ES	1.00												
TIPI-E	.30**	1.00											
TIPI-O	.48**	.39**	1.00										
TIPI-A	.11	-.14	.02	1.00									
TIPI-C	.44**	.13	.36**	.18	1.00								
EPQ-N	-.70**	-.30**	-.37**	-.12	-.31**	1.00							
EPQ-E	.27	.71**	.39**	-.10	.03	-.28*	1.00						
EPQ-P	.07	.26*	.30**	-.11	-.24	-.06	.18	1.00					
NEO-N	-.71**	-.36**	-.39**	-.08	-.41**	.81**	-.26*	.02	1.00				
NEO-E	.44**	.62**	.41**	.03	.29**	-.46**	.74**	.10	-.51**	1.00			
NEO-O	.13	-.13	.42**	.18	.10	.01	-.04	.12	.01	-.06	1.00		
NEO-A	.14	-.18	-.07	.57**	.19	-.30**	-.09	-.35**	-.31**	.09	.10	1.00	
NEO-C	.35**	.10	.15	.11	.72**	-.28*	.03	-.20	-.45**	.28*	.03	.26*	1.00

(- all figures are corrected to two decimal places)

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

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

TIPI – Ten Item Personality Inventory, EPQ – Eysenck's Personality Questionnaire, NEO – NEO Five Factor Inventory.

N- Neuroticism, E – Extraversion, O – Openness, A – Agreeableness, C – Conscientiousness, ES – Emotional Stability, P - Psychoticism

Correlations split by sex and age group:

The data file was split by sex and age group and the correlations repeated, to investigate any variation in rank order between inventories and across scales.

The correlation pattern and figures were almost repeated exactly when compared to the whole sample, for the females and 18 - 35 year age group. The males maintained the significant convergent correlation pattern, with a slight increase in the Neuroticism scale between the NEO-FFI and the EPQ-R short measure, and in the Openness scale between the TIPI and the NEO-FFI; all remaining correlations were lower. The 36 – 59 year age group maintained the significant convergent correlation pattern, with an increase in the correlation value between the NEO-FFI and EPQ-R short measure Neuroticism scales and the TIPI and NEO-FFI Conscientiousness and Openness scales; all remaining correlations were lower. The 60 – 100 year age group failed to maintain the significant convergent correlation pattern, with the sample size being so small the critical value to achieve significance was increased.

Table 6

Rank order comparison of correlations between inventories and across scales, and sub-sample mean correlations between the TIPI and the NEO-FFI.

		Whole sample (N = 81)	Females (n = 60)	Males (n = 21)	18 – 35 yrs (n = 51)	36 – 59 yrs (n = 24)	60 – 100 yrs (n = 6)
TIPI-NEO	**Neuroticism**	1=	2	1	1	3	1
	Extraversion	3	3	3	2	4=	4
	Openness	5	5	2	5	2	5
	Agreeableness	4	4	5	4	4=	3
	Conscientiousness	1=	1	4	3	1	2
TIPI-EPQ	**Neuroticism**	2	2	1	2	1	1
	Extraversion	1	1	2	1	2	2
EPQ-NEO	**Neuroticism**	1	1=	1	1	1	1
	Extraversion	2	1=	2	2	2	2
Mean correlations between TIPI and NEO (r_s)		.61	.63	.55	.59	.65	.62

TIPI – Ten Item Personality Inventory NEO – NEO Five Factor Inventory

EPQ – Eysenck's Personality Questionnaire Revised – short measure

The ranking pattern (see Table 6) was examined across the whole sample and all sub-samples, comparing all three inventories. The whole sample compared closely with the female sub-sample when the TIPI and NEO-FFI were compared, but the remaining four sub-samples failed to follow a similar pattern. The relationship between comparable scales of the NEO-FFI and the EPQ-R short measure remained constant across all samples. The relationship between the comparable scales of the TIPI and the EPQ-R short measure was split, with three of the samples producing a prime ranking position for Extraversion (whole sample, females and 18 – 35 year age group) and the remaining three samples producing a prime ranking position for Neuroticism (Emotional Stability). Criterion validity was assessed by examining the mean correlations between the TIPI and the NEO-FFI for all sub-samples (see Table 6) and compared to those found for the whole sample and in the Furnham (2008) study. All mean correlations were above that identified by Furnham r = .53; and three of the sub-samples were an improvement on that found in the

whole sample r_s = .61: females r_s = .63, 36 – 59 years r_s = .65 and 60 – 100 years r_s = .62 (although non-significant).

P or A and C? Correlations across measures and age and sex differences.

The relationship between the EPQ-R short measure Psychoticism scale and the combination of the Agreeableness and Conscientiousness scales from the TIPI and NEO-FFI measures was examined using correlations. Data were found to be non-parametric therefore correlations were analysed using Spearman's Rho. Correlations were computed for the sample as a whole (see Table 7), the sample split by sex (see Table 8), the sample split by age group (see Table 9) and the sample split by sex and age group (see Table 10).

Table 7

Spearman's Rho Correlations, between EPQ-P and TIPI-P and NEO-P, across the whole sample (N = 81)

	NEO-P	TIPI-P	EPQ-P
NEO-P	1.00		
TIPI-P	.67**	1.00	
EPQ-P	-.36**	-.26*	1.00

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

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

EPQ-P – Eysenck's Personality Questionnaire Psychoticism scale, TIPI-P – Ten-Item Personality Inventory Agreeableness and Conscientiousness scales, NEO-P – NEO Five Factor Inventory Agreeableness and Conscientiousness scales.

There was a significant negative correlation between the EPQ-P and the NEO-P and TIPI-P in the sample as a whole (see Table 7), although significant, the correlations were not strong.

Table 8

Spearman's Rho Correlations, between EPQ-P and TIPI-P and NEO-P, across the sample split by sex.

		NEO-P	TIPI-P	EPQ-P
Female (n = 60)	**NEO-P**	1.00		
	TIPI-P	.74**	1.00	
	EPQ-P	-.44**	-.38**	1.00
Male (n = 21)	**NEO-P**	1.00		
	TIPI-P	.39	1.00	
	EPQ-P	-.07	.03	1.00

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

EPQ-P – Eysenck's Personality Questionnaire Psychoticism scale, TIPI-P – Ten-Item Personality Inventory Agreeableness and Conscientiousness scales, NEO-P – NEO Five Factor Inventory Agreeableness and Conscientiousness scales.

When the sample was split by sex (see Table 8) there continued to be a significant negative correlation between the EPQ-P and the NEO-P and TIPI-P for the females, stronger than that produced by the sample as a whole. However, there was no correlation within the male sub-sample.

Table 9

Spearman's Rho Correlations, between EPQ-P and TIPI-P and NEO-P, across the sample split by age group.

		NEO-P	TIPI-P	EPQ-P
18-35 years (n = 51)	**NEO-P**	1.00		
	TIPI-P	.60**	1.00	
	EPQ-P	-.34*	-.18	1.00
36-59 years (n = 24)	**NEO-P**	1.00		
	TIPI-P	.72**	1.00	
	EPQ-P	-.35	-.33	1.00
60-100 years (n = 6)	**NEO-P**	1.00		
	TIPI-P	.77	1.00	
	EPQ-P	.36	.63	1.00

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

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

EPQ-P – Eysenck's Personality Questionnaire Psychoticism scale, TIPI-P – Ten-Item Personality Inventory Agreeableness and Conscientiousness scales, NEO-P – NEO Five Factor Inventory Agreeableness and Conscientiousness scales.

When the sample was split by age group (see Table 9) the significant negative correlation remained for the 18-35 year age group between the EPQ-P and the NEO-P, but was no longer demonstrated between the EPQ-P and the TIPI-P.

Table 10

Spearman's Rho Correlations, between EPQ-P and TIPI-P and NEO-P, across the sample split by sex and age group.

			NEO-P	TIPI-P	EPQ-P
Female	**18-35 years**	**NEO-P**	1.00		
	(n = 39)	**TIPI-P**	.67**	1.00	
		EPQ-P	-.42**	-.32*	1.00
	36-59 years	**NEO-P**	1.00		
	(n = 18)	**TIPI-P**	.72**	1.00	
		EPQ-P	-.32	-.45	1.00
	60-100 years	**NEO-P**	1.00		
	(n = 3)	**TIPI-P**	.87	1.00	
		EPQ-P	.00	.50	1.00
Male	**18-35 years**	**NEO-P**	1.00		
	(n = 12)	**TIPI-P**	.09	1.00	
		EPQ-P	-.12	.19	1.00
	36-59 years	**NEO-P**	1.00		
	(n = 6)	**TIPI-P**	.52	1.00	
		EPQ-P	-.22	.06	1.00
	60-100 years	**NEO-P**	1.00		
	(n = 3)	**TIPI-P**	.87	1.00	
		EPQ-P	.87	.50	1.00

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

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

EPQ-P – Eysenck's Personality Questionnaire Psychoticism scale, TIPI-P – Ten-Item Personality Inventory Agreeableness and Conscientiousness scales, NEO-P – NEO Five Factor Inventory Agreeableness and Conscientiousness scales.

The significant negative correlation was maintained for the females aged 18-35 years, when the sample was split by sex and age (see Table 10). The remaining correlations between EPQ-P and NEO-P and TIPI-P were non-significant.

Discussion:

The present study primarily aimed to investigate the reliability and validity of the Ten-Item Personality Inventory (Gosling *et al*., 2003) within a British sample, when compared with two longer, more established personality inventories, namely the NEO Five Factor Inventory (Costa & McCrae, 1992a) and the Eysenck Personality Questionnaire-Revised short measure (Eysenck *et al*., 1985). The advantages of a short inventory would be negated if the inventory is not reliably measuring personality traits in-line with longer measures. The results of the current study would appear to support the reliability and validity of the TIPI for use within a British population.

Responses to items on all three inventories are on different scales; the TIPI uses a five point Likert scale, the NEO-FFI uses a seven point Likert scale and the EPQ-R short measure uses yes/ no. The resultant means (SD) were therefore not directly comparable, subsequently a ranking system was utilised to compare the descriptive statistics across inventories. Analysis of the ranking patterns provided little insight

into the relationships between inventories. The most notable pattern to emerge was the high ranking positions of Agreeableness and Conscientiousness across measures, between sexes and across age groups, and the low ranking of Psychoticism, which by some (e.g. Eysenck, 1992) is considered to be composed of Agreeableness and Conscientiousness reversed. The related ranking positions of these three scales may be interpreted as indicating measures of the same trait, although more detailed analysis would be warranted.

The EPQ-R short measure was the only inventory to maintain the same ranking pattern across the whole sample and all sub-samples. This may be indicative of the most stable inventory across sex and age, or it may be the result of the measure of three traits as opposed to the five traits measured by the other two inventories, thereby increasing the range of possible ranking patterns.

Gosling *et al*. (2003) advise caution regarding factor analysis not being the most pertinent method of analysing the psychometric properties of the TIPI. They recommend test-retest reliability as the method of choice when examining the reliability of the TIPI. However, they do recognise the value of ongoing psychometric testing within different samples in order to build its psychometric profile. They state their concern was primarily about the validity of the TIPI, rather than in its reliability.

Two item scales are considered to be renowned for their poor performance when analysed for alphas and factor structure (Kline, 2000). The results of the principal component analysis, with varimax rotation, of the TIPI supported the claim that a poor factor structure would result when analysing a two item scale. The highest factor loadings were examined for each of the ten items; the only factor structure to support the reliability of the TIPI was factor 1, indicating that Extraversion is a valid measure within the population sampled. The remaining four components did not support the reliability of the TIPI on the remaining four personality scales. Test-retest may have been a more pertinent test to use to assess the reliability within the current population; however time restraints were a consideration. The limitation of factor analysis providing only descriptive, and not explanatory, worth to trait personality theory should also be recognised (Block, 2001).

The reliability indices, all .7 or above, using Cronbach's alpha, supported the high internal reliability of the NEO-FFI and the EPQ-R short measure, with the exception of the Psychoticism scale on the EPQ-R short measure. The alpha of .4 for the Psychoticism scale is considered to be poor (Field, 2005). The internal reliability of the Psychoticism scale has proved problematic and several alterations have been made over time in an attempt to improve it: Eysenck *et al*. (1985) does note that the adjustments have improved the distribution and increased the mean scores. However, this does not hold true for the population in the current study. Since the initial development of the Psychoticism scale was undertaken within a psychiatric setting, this may lead to a difference in results when compared with another sample. The improved scale was, however, tested on the general population, thus a higher level of internal reliability may have been expected within the current study. The current sample was small, and it has been noted that individuals who may score highly on the Psychoticism scale are more reluctant to participate in research (Eysenck *et al*., 1985), both possible explanations for the low level of internal reliability found. Furnham *et al*. (2008) does state that further development of the Psychoticism scale is necessary to improve its reliability.

The internal reliability of the TIPI varied across the scales, the Emotional Stability and Extraversion scales demonstrating high internal reliabilities, with alphas of .7 or above. The remaining three scales demonstrated poor internal reliabilities, .5 or below, commensurate with the warning given by Gosling *et al*. (2003). The comparison of reliability indices of the three inventories with those reported in the literature regarding the development of the inventories provided mixed results. The EPQ-R short measure alphas compared exactly with those reported by Eysenck *et al*. (1985) for the Neuroticism and Extraversion scales, supporting the reliability of the EPQ-R short measure in the population sampled. However, the Psychoticism alpha was markedly lower as previously discussed.

The comparison of the NEO-FFI alphas with those reported by Costa and McCrae (1992a), suggested that the reliability was increased slightly within the current population, especially for the Agreeableness and Conscientiousness scales. The female domination of the sample may have resulted in the higher alphas for the Agreeableness and Conscientiousness scales, as females are reported to score higher on these traits (Feingold, 1994). The TIPI alphas compared closely with the original alphas reported by Gosling *et al*. (2003) for the Emotional Stability, Extraversion, Openness and Agreeableness scales, but the Conscientiousness alpha was markedly lower in the current study, .36 compared to .50 in the original study. Both alphas are indicative of poor reliability within the populations sampled. The low coefficient alphas explaining the reliability indices of the TIPI provide further support for Gosling's warning that alphas are not a dependable indicator of the reliability of a measure consisting of two items per scale. However, it may be concluded that the Emotional Stability and Extraversion scales do provide a reliable measure within the population sampled.

The data were found to be non-parametric, therefore correlations were analysed using Spearman's Rho. The convergent correlations across inventories were as expected and may be interpreted as the best indicator of the validity of the TIPI in a British sample when compared to the NEO-FFI and the EPQ-R short measure. The mean convergent correlation between the TIPI and the NEO-FFI is r_s = .61, which is an improvement on that found by Furnham (2008) r = .53, although Furnham chose to compare the median convergent correlations, which showed an improved figure of r = .58. The current sample was spread further across the population as a whole, as opposed to Furnham's study, which focussed on a student population (mean age of 19.7 years, SD = 2.1). However, the present study did simulate the Furnham study as far as sex distribution was concerned, with a larger percentage of female participants.

When the German version of the TIPI was compared to the NEO-FFI the mean convergent correlation was r = .38 (Herzberg & Brähler, 2006), indicating that trials within specific populations are necessary prior to generalisations regarding the validity of the TIPI as a reliable short measure of the Big Five personality traits. The results of the current study need to be interpreted with reference to the population sampled.

The individual convergent correlations varied in strength, the strongest being for the Emotional Stability (Neuroticism reversed) scale and the weakest being for the Openness scale. The Emotional Stability (Neuroticism reversed) and Extraversion correlations across all three inventories demonstrated strong correlations, and the

Conscientiousness correlation between the two inventories also demonstrated a strong correlation, indicating that the TIPI is most robust when measuring these three traits as opposed to Openness and Agreeableness. This rank-order pattern does not entirely support that found in the Furnham (2008) study, with the Extraversion correlation lower and the Openness correlation higher. The different population sampled may explain the inconsistencies found.

The age at which personality traits are considered to be stable is debatable, Terracciano *et al*. (2006) state this is so after the age of thirty, but Allemand *et al*. (2008) propose that personality traits continue to fluctuate in to old age. This may explain the variation in trait rank-order found between the current study and the Furnham (2008) study. It should also be noted that the correlations between the NEO-FFI and EPQ-R short measure Neuroticism and Extraversion scales were significant, positive and strong.

Anomalies found within the correlation table were subjected to scrutiny, the TIPI Openness scale significantly correlated with seven out of the ten possible scales, at about the level of .4, either negatively or positively. Openness is considered by some (e.g. Peabody & Goldberg, 1989) to be more a measure of intellect than personality. Since the present results indicate its relationship with so many scales this may be indicative that it is not a measure of a single personality trait but one which overlaps more generally. Although it should be noted that the NEO-FFI Openness scale only correlated significantly with the TIPI Openness scale, all remaining correlations were negligible. Therefore, it may be surmised that the TIPI scale is not a good measure of the Openness trait. The TIPI Agreeableness scale correlated significantly with only one scale, the NEO-FFI Agreeableness scale, demonstrating it to be a measure of Agreeableness alone.

Correlations were repeated for the sub-samples and compared to those reported for the sample as a whole. The mean correlations for the sub-samples between the TIPI and the NEO-FFI were all higher than those found in the Furnham (2008) study; indicating the TIPI as a valid measure of the Big-Five personality domains, across British age groups, when compared with the NEO-FFI.

There is ongoing debate regarding the relationship between the Psychoticism scale and the Agreeableness and Conscientiousness scales. Eysenck (1992) stating that Agreeableness and Conscientiousness are sub-scales of the super trait Psychoticism, and Costa and McCrae (1992b) stating that they are independent high order traits. To investigate this in the current sample the scores for the Agreeableness and Conscientiousness scales on the TIPI and NEO-FFI inventories were combined for each participant, and referred to as TIPI-P and NEO-P, and were correlated with the Psychoticism score on the EPQ-R short measure. The sample was split according to sex and age and further correlations were computed. When the sample is considered as a whole, significant negative correlations were found between EPQ-P (EPQ-R short measure Psychoticism scale) and the TIPI-P and NEO-P. Although the correlation was not strong, it was indicative that the scales were measuring aspects of a similar trait, lending support to Eysenck's stance. When the sample is split by sex the trend continues for females but not for males. There is a stronger significant negative correlation when only females are considered.

When the sample is split by age group only the youngest age group maintains the relationship between the EPQ-P and the NEO-P, the TIPI-P is no longer significantly negatively correlated. When the sample is split by sex and age the trend continues for the youngest female group only. The middle aged female group maintains a negative correlation of below -.3, but it fails to be significant, again possibly due to the increase in the critical value due to the small sample size. The remaining correlations are not indicative. The results thus support Eysenck's theory when the sample as a whole and the young female group is considered. However, in all other correlations the scales may be measuring different personality traits, possibly supporting Costa and McCrae. It should be noted that within the current sample the Psychoticism scale had poor internal reliability, which may have had an effect on the subsequent correlations.

Limitations:

Recognition of the specific limitations of research design when adopting questionnaire methodology need to be considered when drawing conclusions from the results of the current study. The self-report aspect of questionnaires relies heavily on individuals evaluations when completing the measure. Within the current design an on-line recruitment method was utilised, thus it is not possible to be confident that all individuals completing the inventories met the research criteria and responded accurately. An assessment of the costs and benefits of such a recruitment method need to be closely analysed to ascertain the design choices. Within a small research project such as this, with strict time constraints, it was felt to be the optimal method of recruitment in order to access a broad sample of the population. If the research were to be conducted on a narrower population, alternative recruitment methods may have been preferential. It should be noted that this chosen method of recruitment discriminates against individuals who are not computer literate. A mixed method of recruitment may be the most appropriate option to access a more representative sample. The faking of personality inventories has been subjected to research (e.g. Martin *et al*., 2002) and ipsative methods are preferred as individuals have been found to provide fake responses less frequently than in normative methods.

Low response rates are common in independent self-report questionnaire designs; although within the current study the response rate was impressive considering the number of individuals initially approached. The number of participants overall did not exceed one hundred, small when compared to the sample size of similar studies, with the exception of Furnham (2008). It should be noted that an increase in sample size does not guarantee a significant result, but in the current study a larger sample may have provided more equal group sizes considering sex and age differences.

Completion of personality inventories may be considered relatively low risk for affecting psychological well-being. Whilst contact details of relevant potentially helpful organisations were provided personal follow-up was not possible with the recruitment method chosen. It was therefore not possible to ascertain whether there were any adverse impacts of completing the questionnaires. Ethical standards were adhered to, but this issue may be an important consideration with increasing use of cyber technology for research. A more robust support system for participants involved in online research may need to be more generally considered.

Conclusion:

Psychometric assessment of personality inventories is an ongoing process, ensuring they remain reliable and valid for the population in which they are being utilised. Accumulation of psychometric support lends confidence to the widespread use of an inventory. The TIPI is a convenient method of assessing the Big-Five personality traits, taking only one minute to complete, when time is a consideration or when personality is not the prime focus of the research. The evidence from the current study provides an indication that the TIPI may be a reliable and valid inventory for use within a British population, when circumstances warrant it, supporting the findings of Furnham (2008). An inventory this concise should not eclipse the utilisation of more detailed inventories; however, it does have a place in the psychometric testing of personality traits. Further investigation of the TIPI's psychometric properties within different populations is considered to be warranted.

References:

Alexopoulos, D.S. & Kalaitzidis, I. (2004). Psychometric properties of Eysenck Personality Questionnaire-Revised (EPQ-R) short scale in Greece. *Personality and Individual Differences*, 37, 1205-1220.

Allemand, M., Zimprich, D. & Hendriks, A.A.J. (2008). Age differences in five personality domains across the life span. *Developmental Psychology, 44*(3), 758-770.

Allport, G.W. (1937). *Personality: A psychological interpretation*. London: Constable.

Allport, G.W. & Odbert, H.S. (1936). Traitnames: A psycho-lexical study. *Psychological Monographs, 47*, 1-171.

Aluja, A., Garcia, O. & Garcia, L.F. (2003). A psychometric analysis of the revised Eysenck Personality Questionnaire short scale. *Personality and Individual Differences, 35*, 449-460.

Barelds, D.P.H. & Luteijn, F. (2002). Measuring personality: a comparison of three personality questionnaires in the Netherlands. *Personality and Individual Differences, 33*, 499-510.

Benet-Martínez, V. & John, O.P. (1998). Los cinco grandes across cultures and ethnic groups: Multitrait multimethod analyses of the big five in Spanish and English. *Journal of Personality and Social Psychology, 75*(3), 729-750.

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, D.M. & Craik, K.H. (1983). Act prediction and the conceptual analysis of personality scales: Indices of act density, bipolarity, and extensity. *Journal of Personality and Social Psychology, 45*(5), 1081-1095.

Cattell, R.B. (1957). *Personality and motivation structure and measurement*. Oxford: World Book.

Cattell, R.B., Cattell, A.K. & Cattell, H.E.P. (1993). *Sixteen personality factor questionnaire, fifth edition.* Champaign, Illinois: Institute for Personality and Ability Testing.

Coolican, H. (2006). *Introduction to research methods in psychology* (3rd edn). London: Hodder Arnold.

Costa, P.T. & McCrae, R.R. (1985). *The NEO Personality Inventory*. Odessa, Florida: Psychological Assessment Resources.

Costa, P.T. & McCrae, R.R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology, 54*(5), 853-863.

Costa, P.T. & McCrae, R.R. (1992a). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, Florida: Psychological Assessment Resources.

Costa, P.T. & McCrae, R.R. (1992b). Four ways five factors are basic. *Personality and Individual Differences, 13*(6), 653-665.

Costa, P.T., Terracciano, A. & McCrae, R.R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology, 81*(2), 322-331.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.

Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417-440.

Ehrhart, M.G., Ehrhart, K.H., Roesch, S.C., Chung-Herrera, B.G., Nadler, K. & Bradshaw, K. (2009). Testing the latent factor structure and construct validity of the Ten-Item Personality Inventory. *Personality and Individual Differences, 47*, 900-905.

Erikson, E.H. (1966). Eight ages of man. *International Journal of Psychiatry, 2*(3), 281-300.

Eysenck, H.J. (1947). *Dimensions of personality*. London: Kegan Paul.

Eysenck, H.J. (1957). *The dynamics of anxiety and hysteria*. London: Routledge and Kegan Paul.

Eysenck, H.J. (1967). *The biological basis of personality*. Springfield, Illinois: Thomas.

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. (1992). Four ways five factors are not basic. *Personality and Individual Differences, 13*(6), 667-673.

Eysenck, H.J. & Eysenck, S.B.G. (1964). *Manual of the Eysenck Personality Inventory*. London: University Press.

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

Eysenck, H.J. & Eysenck, M.W. (1985). *Personality and individual differences*. New York: Plenum.

Eysenck, S.B.G., Eysenck, H.J. & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences, 6*(1), 21-29.

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116*(3), 429-456.

Field, A. (2005). *Discovering statistics using SPSS*. London: Sage.

Francis, L.J., Lewis, C.A. & Ziebertz, H. (2006). The short-form revised Eysenck Personality Questionnaire (EPQR-S): A German edition. *Social Behavior and Personality, 34*(2), 197-204.

Furnham, A. (2008). Relationship between four big five measures of different lengths. *Psychological Reports, 102*, 312-316.

Furnham, A., Eysenck, S.B.G. & Saklofske, D.H. (2008). The Eysenck personality measures: Fifty years of scale development. In G.J. Boyle, G. Matthews & D.H. Saklofske (eds) *The sage handbook of personality theory and assessment, vol. 2: Personality measurement and testing* p.199-218. London: Sage.

Furnham, A. & Chamorro-Premuzic, T. (2004). Estimating one's own personality and intelligence scores. *British Journal of Psychology, 95*, 149-160.

Gosling, S.D., Rentfrow, P.J. & Swann, W.B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*, 504-528.

Guilford, J.P. & Guilford, R.B. (1934). An analysis of the factors in a typical test of introversion-extroversion. *The Journal of Abnormal and Social Psychology, 28*(4), 377-399.

Harre, R. & Gillett, G. (1994). *The discursive mind*. London: Sage.

Hathaway, S.R. & McKinley, J.C. (1967). *Minnesota multiphasic personality inventory manual*. New York: Psychological Corporation.

Herzberg, P.Y. & Brähler, E. (2006). Assessing the big-five personality domains via short forms: A cautionary note and a proposal. *European Journal of Psychological Assessment, 22*(3), 139-148.

Hofmans, J., Kuppens, P. & Allik, J. (2008). Is short in length short in content? An examination of the domain representation of the Ten Item Personality Inventory scales in Dutch language. *Personality and Individual Differences, 45*, 750-755.

Jolliffe, I.T. (1972). Discarding variables in a principal component analysis, I: artificial data. *Applied Statistics, 21*, 160-173.

Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 133-145.

Keirsey, D. (1998). *Please understand me II: Temperament, character, intelligence.* California: Prometheus Nemesis Book Company.

Kline, P. (2000). *The handbook of psychological testing* (2nd edn). London: Routledge.

Larsen, R.J. & Buss, D.M. (2005). *Personality psychology: Domains of knowledge about human nature* (2nd edn). New York: McGraw-Hill.

Martin, B.A., Bowen, C.C. & Hunt, S.T. (2002). How effective are people at faking personality questionnaires? *Personality and Individual Differences, 32*, 247-256.

Matthews, G., Deary, I.J. & Whiteman, M.C. (2003). *Personality traits* (2nd edn). Cambridge: University Press.

McAdams, D.P. (1992). The five-factor model in personality: A critical appraisal. *Journal of Personality, 60*(2), 329-361.

McAdams, D.P. (2009). *The person: An introduction to the science of personality psychology* (5th edn). New Jersey: Wiley.

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

McCrae, R.R. & Costa, P.T. (1995). Trait explanations in personality psychology. *European Journal of Personality, 9*(4), 231-252.

McCrae, R.R., Costa, P.T., de Lima, M.P., Simões, A., Ostendorf, F., Angleitner, A. *et al*. (1999). Age differences in personality across the adult life-span: Parallels in five cultures. *Developmental Psychology, 35*(2), 466-477.

Mischel, W. (1996). *Personality and assessment*. New Jersey: Lawrence Erlbaum.

Mischel, W., Shoda, Y. & Smith, R.E. (2003). *Introduction to personality: Toward an integration* (7th edn). New Jersey: Wiley.

Muck, P.M., Hell, B. & Gosling, S.D. (2007). Construct validation of a short five-factor model instrument: A self-peer study on the German adaptation of the Ten-Item Personality Inventory (TIPI). *European Journal of Psychological Assessment, 23*(3), 166-175.

Myers, I.B., McCaulley, M.H., Quenk, N.L. & Hammer, A.L. (1998). *Manual: A guide to the development and use of the Myers-Briggs type indicator*. Paolo Alto: Consulting Psychologists Press.

Norman, W.T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66*(6), 574-583.

Peabody, D. & Goldberg, L.R. (1989). Some determinants of factor structures from personality-trait descriptors. *Journal of Personality and Social Psychology, 57*(3), 552-567.

Stelmack, R.M. & Stalikas, A. (1991). Galen and the humour theory of temperament. *Personality and Individual Differences, 12*(3), 255-263.

Terracciano, A., Costa, P.T. & McCrae, R.R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin, 32*(8), 999-1009.

Woods, S.A. & Hampson, S.E. (2005). Measuring the big five with single items using a bipolar response scale. *European Journal of Personality, 19*(5), 373-390.

Zuckerman, M., Kuhlman, D.M., Joireman, J. Teta, P. & Kraft, M. (1993). A comparison of three structural models for personality: The Big Three, the Big Five, and the Alternative Five. *Journal of Personality and Social Psychology, 65*(4), 757-768.