The Stability of the Five-Factor Model of Personality in Personnel Selection and Assessment in Greece

Ioannis Tsaousis* and Ioannis E. Nikolaou

The stability and replicability of the Five-Factor model of personality across samples and testing purposes remain a significant issue in personnel selection and assessment. The present study explores the stability of a new Greek Big Five personality measure (TPQue) across different samples in order to explore the suitability of the measure in personnel selection and assessment. The factor structure of the measure across three samples (students, employees, and job applicants) is examined. The results of exploratory and confirmatory factor analyses show that the five-factor structure remains intact for the students', the applicants' and the employees' samples – contrary to previous studies – with all the sub-scales of the personality measure (TPQue) loading on the intended factors. Furthermore, congruence coefficients between the samples justify the stability of the model in the working settings.

The emergence of the five-factor model of personality (FFM) has provided personnel researchers and practitioners with a very useful taxonomy for examining the application and use of personality dispositions in work settings. As Schmit and Ryan (1993) noted, most studies examining the construct-validity and factor structure of the FFM have made use of volunteer samples, rather than employees or job applicants, and therefore, similar factor structures should not be assumed across different samples. The scope of the present study is to examine the factorial stability of a new Greek Big Five measure (Tsaousis 1999) in two occupational groups (i.e. employees, and job applicants), in order to establish whether the five-factor structure remains intact across occupational samples and testing purposes.

The Big Five at Work

The development of a well-accepted taxonomy of personality traits has occurred as a result of the independent work of a number of researchers since the early 1960s (for a review see Digman 1990; John 1990; Matthews 1997; McCrae and John 1992). Although some disagreement exists regarding the precise factor elements and the naming of the five factors, they are usually labelled as follows: extraversion or surgency, agreeableness, conscientiousness, neuroticism or emotional stability, and openness to experience or intellect. They have been referred as 'robust' across samples, rating formats, methods of factor analysis, cultures and languages. Particularly, cross-cultural studies have shown that the five-factor model emerges in various languages and different socio-economic environments, providing support for the possibility that those dimensions might represent cultural 'universals' of person description (Goldberg 1981). More specifically, apart from the American-English factor structure, for which enormous amounts of data are available (Costa and McCrae 1992; Goldberg 1990), there are also German, Dutch, Belgian, Czech, Polish, Russian, Italian, Spanish, Greek, Hebrew, Hungarian, Estonian, Finnish, Turkish, Korean, Filipino, Chinese, and Japanese replications of the five factor structure following either etic or emic procedures (Berry 1969). For an excellent review in cross-cultural research as well as for a comparison between lexical studies among languages, see Saucier *et al*. (in press).

However, this theoretical framework is not without such criticisms as: the model is not comprehensive, it is method-bound, it confounds other important constructs and that its level of measurement is very broad (Hough 1997; Matthews 1997). Despite these criticisms, however, the FFM has provided personality psychology with a clear measurement framework and is responsible for the resurgence of interest of personality in the field of work and organizational psychology.

A number of meta-analytic studies have investigated the validity of the FFM across different types of occupations (Barrick and Mount 1991; Tett *et al*. 1991; Salgado 1997), such as in jobs involving teamwork and

Address for Correspondence: Ioannis Tsaousis, Department of Sciences in Pre-School Education, University of the Aegean, Leof. Dimocratias 1, Rhodes, 85 100, Greece. E-mail: tsaousis@rhodes.aegean.gr

interpersonal interaction (Mount *et al*. 1998) as well as with army samples (Salgado 1998). The findings of these studies suggest that a 'conscientiousness-plus' factor (Mount and Barrick 1995) is a strong predictor of overall job performance across occupations and hierarchical positions. This factor does not encompass elements only from conscientiousness, but also from agreeableness, emotional stability and extraversion. Mount and Barrick (1995) called this factor *'functional personality at work'*, whereas Schmit and Ryan (1993), using a job applicants sample, called it *'ideal employee factor'*. However, Robertson and Callinan (1998) argue that a careful examination of the uncorrected validity coefficients of these studies suggests that very little has changed since the early 1970s and the pessimistic review of Ghiselli (1973) on the validity of personality testing. The upper limit of personality traits against overall job performance rarely exceeds .40 after meta-analytic corrections.

Most of these studies, as Robertson and Callinan (1998) noted, have used personality variables similar, in a way, to early research in the field of general mental ability, where a unidimensional construct was linearly correlated with measures of overall job performance. Personality, however, differs from mental ability. Although linear relationships are expected between personality dimensions of a model such as the FFM, which is a prerequisite for the all the factor analyses carried out identifying the Big-Five structure, curvilinear relationships should be expected between specific personality constructs and job performance. Employees with very high scores on specific aspects of conscientiousness, for example, order or deliberation and self-discipline may lack the flexibility to adapt when circumstances suddenly change. Second, strong moderating effects should also be expected between personality and work-related outcomes such as the moderating effect of job type, organizational-environmental characteristics, task-goal characteristics, etc. (Schneider and Hough 1995).

Barrick and Mount (1993) showed that the degree of autonomy moderates the criterion-related validity of some personality dimensions of the FFM in a sample of managers. The validity of conscientiousness, extraversion and agreeableness was greater for managers in jobs high in autonomy compared with those in jobs low in autonomy. The correlation was positive for the first two and negative for the third personality dimension. In another study, Barrick, Mount, and Strauss (1993) demonstrated that autonomous goal setting and, to a lesser extent, goal commitment mediate the predictive validity of conscientiousness. Using structural equation modelling the authors showed that sales representatives high in conscientiousness are more likely to set goals and be more committed to them, which is then positively associated with sales volume and supervisory ratings of job performance, compared to sales representatives low in conscientiousness. Finally, Nikolaou and Robertson (in press) found that the type of job (i.e., jobs involving or not interaction with others) moderates the criterion-related validity of agreeableness, extraversion, and openness to experience. Agreeableness and Openness to Experience were positively related to overall job performance for employees working in occupations, involving interpersonal interaction, such as teachers, managers, and sales representatives, compared to employees working in posts involving less interpersonal interaction; extraversion was also negatively related to job performance for this type of occupations.

Another issue which has recently attracted more attention is the appropriate level of analysis in personality research, as expressed in the 'bandwidth-fidelity' dilemma (Hough and Oswald 2000). Strong arguments are raised in favour of each approach; the supporters of high bandwidth cite the supportive results of studies examining the predictive validity of integrity or employee reliability tests (Hogan and Hogan 1989), and service orientation. Integrity tests measure a composite of conscientiousness, emotional stability and agreeableness that normally elicits correlations in the range of .40 for predicting supervisory ratings of job performance (Ones *et al*. 1996). Service orientation (Hogan *et al*. 1984) is another broad personality construct composed of agreeableness, conscientiousness, and emotional stability.

These two broad-band constructs better predict wider work-related outcomes, such as overall job performance, contrary to narrow dependent variables, such as work competencies. On the other hand, the advocates of the low fidelity approach support the importance in following a confirmatory approach when assigning relevant personality dimensions to appropriate performance criteria. They argue that the choice of the appropriate level of analysis in the personality domain and the selection of the most relevant criteria should result from either an empirically driven search of the literature or preferably as the result of a job analysis.

Although it is impossible to determine work behaviour from the effect of personality alone – because of the importance of situational variables such as colleagues, supervision, job environments, reward structure etc. (Robertson and Callinan 1998) – personality variables can be significant predictors of work performance when

they are carefully matched with appropriate occupations and organizations. Day and Silverman (1989) used a small sample of accountants and showed that when a preceding job analysis has identified potentially useful personality constructs, these constructs can explain incremental variance for a standard cognitive ability test. Schmidt and Hunter (1998), in a paper reviewing the validity of 19 selection procedures and their incremental validity when paired with measures of general mental ability, showed that a combination of integrity tests with cognitive ability measures produced the highest multiple validity coefficients (multiple $r = .65 - .67$) than that of any other combination, regarding the criteria of overall job performance and performance in job training programmes.

The Stability of Big Five in Selection and Assessment

The first researchers that questioned the stability of the five-factor model across different populations were Schmit and Ryan (1993). They responded to the inconsistent results on the validity of the five-factor model used in occupational settings by arguing that these may have resulted from the almost exclusive use of volunteer samples; they claimed that people would be expected to respond differently in settings where 'strong' situational effects exist, such as in recruitment or performance appraisal. It is highly likely, therefore, that job applicants or employees were guided by self-presentation schemas, wishing to show an image of competence, or of fulfilling the job requirements. Schmit and Ryan (1993) challenged the adequacy of the five-factor model when the above conditions persist. They examined their hypothesis by comparing the factor structure of a well-established measure of the five-factor model (NEO-FFI; Costa and McCrae 1989) across two samples; a sample of university students and a sample of job applicants 'seeking employment assistance at a university' (p. 969).

Their results showed that although the Big Five structure fit well the volunteers-student sample, this was not the case for job applicants. A six-factor solution worked more adequately for the latter sample. The first factor of this solution included loadings from four of the five dimensions of Big-Five, namely Conscientiousness, Extraversion, Agreeableness, and the opposite pole of Neuroticism. Schmit and Ryan (1993) claimed that this factor should be called '*ideal-employee*' factor, describing an applicant as productive, highly likeable, conscientious, courteous, thoughtful, considerate, organized, committed, active, high spirited, not helpless, and not argumentative. These are necessary attributes for a successful candidate suggesting, according to Schmit and Ryan (1993), that job applicants can change their responses in a personality questionnaire, in order to show themselves in a favourable way, and increase their changes in the recruitment process. Nevertheless, Costa (1996) argued that when he re-analysed Schmit and Ryan's (1993) data, requesting a five-factor solution, only Agreeableness was not very clearly identified with the other four factors easily recognized. He concluded that 'the evaluation bias on the structure of the personality questionnaire are relatively modest' (Costa 1996, p. 231).

Two more studies examined the stability of the FFM in applied settings. Montag and Levin (1994) compared the structure of NEO-PI-R (Costa and McCrae 1992) in two applicant samples of males and females. Their results showed that a clear five-factor solution came up for the female sample, but a less clear-cut solution was produced for the male sample, yielding four to five factors moderately resembling to the normative data. However, the results of the factor analysis for the combined sample found a clear five-factor solution highly congruent with the American combined sample. Finally, Cellar *et al*. (1996) in a similar examination using a sample of flight attendant trainees identified better fit for a six-factor solution. Their sixth factor, however, did not resemble Schmit and Ryan's *'ideal-employee'* factor, rather including items from Neuroticism and the opposite pole of Conscientiousness.

The results of these studies call for a careful examination of the stability of Big-Five in applied settings, such as in selection and assessment, and the effects of self-deception and impression-management in personality testing. Although most personality measures usually include some scales to detect intentional distortion, social desirability and faking patterns, experience and research has shown that job applicants have often higher scores on these scales than employess or students (Barrick and Mount 1996). Nevertheless, their effect on the criterion-related validity and on the convergent and discriminant validity of the five-factor model is negligible, as the most recent reviews of the topic have shown (Barrick and Mount 1996; Ones and Viswesvaran 1998; Ones *et al*. 1996) although they may have a dramatic effect on actual hiring decisions (Rosse *et al*. 1998).

Despite the minimal effect of intentional distortion and social desirability on the validity of Big Five personality questionnaires, the question of the stability of the five-factor model in job applicants sample, as opposed to employees or volunteers, remains very significant for the use of Big Five in selection and assessment settings.

The present study examines further the stability of Big Five in two occupational groups of employees and job applicants as opposed to a non-applicant sample. Assumptions made for this study are the following:

1. The psychometric structure of the questionnaire should be the same across the three groups. In other words, the factor structure of the personality measure across the two occupational samples (employees-job applicants) would be similar to the normative sample of the instrument (non-applicants).
2. The factor structure will be defined according to items loaded onto a scale and the relationships and correlations between those scales.

Method

Participants

Three independent samples were used. The first one consisted of the normative sample of the personality measure (TPQue), composed of 1054 university students, 410 (39%) out of whom were males and 644 (61%) were females. The average age of the individuals was 19.9 years old (SD = 4.32). The second sample consisted of 561 job applicants, applying for various openings in a national bank through a recruitment firm. 197 (35.1%) were males and 364 (64.9%) were females; the average mean age was 26.8 years old (SD = 4.35). The third sample consisted of 225 employees; 57% of them were females, and the vast majority (83%) below the age of 40, with a mean tenure of 4.3 years. They were employed in a variety of organizations – of small and medium size mainly – and positions, such as teachers (19.3%), in sales/insurance (18.8%), managers (15.6%), administration (11.5%), in arts/design (10.9%), in personnel/training (8.9%), as accountants (8.3%), and in scientific positions (6.8%). Job applicants had completed the personality measure as part of the recruitment procedure. Employees had filled in the personality questionnaire as part of a research project.

Materials

The main instrument used in this study was a Greek personality questionnaire based on the five-factor model (FFM). The Trait Personality Questionnaire (TPQue) (Tsaousis 1999), which is the only measure of the FFM in Greece, is a comprehensive measure of the five major dimensions or factors of personality as well as of the most important traits that define each domain in the Greek language. It is based on Costa and McCrae's (1992) definitions of the most acceptable factors in the five-factor theory (Extraversion, Neuroticism, Openness to Experience, Agreeableness, and Conscientiousness) taking into account the specific ethnic and cultural characteristics of the Greek population (Tsaousis 1996).

The development of the items was based on both rational and factor analytic techniques. The final version of the TPQue consisted of 180 items measuring the five broad factors and thirty specific sub-scales, which correspond to the most influential traits of the domain as well as 26 items which were targeted as identifying the response sets of lying and social desirability. Each factor consisted of 36 items and each sub-scale consisted of 6 items. Item responses were recorded on a 5-point scale ranging from strongly disagree (1) to strongly agree (5).

Coefficient alpha reliabilities were .88, .89, .83, .78, and .88 for Extraversion, Neuroticism, Openness to Experience, Agreeableness, and Conscientiousness, respectively, while the alpha coefficients for each of the sub-scales ranged between .51 and .80. In addition, the TPQue has indicated an acceptable evidence of concurrent and construct validity (e.g. internal validity-factorial structure, convergent and discriminant validity). For a more thorough description of the questionnaire (item content, development methods, and various psychometric properties), see Tsaousis (1999).

Results

Table 1, along with the descriptive statistics (means and standard deviations) provides also the alpha reliabilities for the five factors across the three samples. As can be seen, the internal consistency of the personality measure across the two occupational samples of employees and job applicants, is very similar to the normative sample (non-applicants).

The analysis with respect to the first research question, which addresses the stability and replicability of the factor structure of the model across three different samples, will be considered in two different ways: First, by using exploratory factor analysis (principal axis method) the factor structure of each of the three different pairs of comparison (i.e., applicants vs. employees, applicants vs. non-applicants, and employees vs. non-applicants) will be examined so as to evaluate the degree of their factorial similarity. The assumption is that the factor structure across the different pairs will be similar. To test the above hypothesis, congruence coefficients (Harman 1976) between the contrasted pairs will be calculated.

Second, the stability of the three different factor structures can also be demonstrated by the

Table 1: Descriptive statistics of the five factors across the three samples

Factors	E	N	O	A	C
Applicants (N = 561)					
Alpha	.88	.90	.82	.78	.89
Mean	132.19	88.07	124.66	122.62	139.39
SD	15.06	17.26	13.96	11.97	15.15
Employees (N = 225)					
Alpha	.89	.92	.85	.79	.90
Mean	128.10	100.42	125.19	122.06	129.76
SD	16.13	19.55	14.91	11.91	16.83
Non-Applicants (N = 1054)					
Alpha	.88	.89	.83	.78	.88
Mean	122.40	112.25	126.48	118.44	115.07
SD	17.16	18.96	15.69	13.07	17.43

matrices of the intercorrelations of the TPQue scores. Fisher's *z* will be calculated and post hoc power test will be applied, to test the significance of the differences between the correlation coefficients of the three contrasted groups.

Finally, using confirmatory factor analysis, we will demonstrate whether the same five factor model holds for each of the three populations. For that reason, absolute goodness-of-fit indices (χ^2/df) as well as relative goodness-of-fit indices, such as the Tucker-Lewis's (TLI), the Normed Fit (NFI), the Normed Noncentrality Fit (CFI), and Root Mean Square Error of Approximation (RMSEA) along with parsimony goodness-of-fit indices (the Parsimonious Normed Fit -PNFI)) will be calculated for each group as an indicator of model fit.

Table 2 presents the factor solutions between the first contrasted pair: applicants vs. employees as well as the coefficients of congruence between factors in the two samples.

Table 2 contains the factor loadings of the TPQue sub-scales on the five factors for both applicant and employee samples. As can be seen, in both samples the TPQue sub-scales load on the intended factor, which indicated the stability of the Five-Factor model in both conditions. Where secondary loadings appeared, they were meaningful and consistent with the theory. For example, in the applicant sample, O4 'Actions' loaded strongly on Extraversion factor along with the Openness Factor (where it actually belongs) since according to the conceptual model proposed by Costa and McCrae (1992) people who like to try new activities and visit new places they are also outgoing and energetic individuals, they have a positive perspective towards life, they look forward to experiencing new adventures, which are characteristics for extraverted individuals. Similarly, in the employee sample, A5 'Modesty' loaded negatively on Extraversion factor along with the Agreeableness factor (where it actually belongs) since humble and self-effacing people are totally different from extraverted individuals (negative loading) who usually like to talk a lot about themselves (Costa and McCrae 1992).

In terms of the amount of the variance explained, we see that in both the applicant sample and in the employee sample the variance explained by the particular solutions is high and similar to each other. The factor structure for the applicant sample accounts for almost 57% of the total variance, while the factor structure for the employee sample accounts for almost 59% of the total variance. Additionally, all the five congruence coefficients were above the critical value of .90 (Harman 1976). More specifically, the coefficients were .95 for E, .96 for N, .97 for O, .95 for A, and .92 for C. Thus, the factor structure seemed to be stable either across applicant or employee samples.

Table 3 presents the factor solutions between the second contrasted pair, (applicants vs. non-applicants) as well as the congruence coefficients between the five-factor scales in the two particular samples.

As we can observe, all the TPQue sub-scales of both samples load on the intended factor, while any secondary loadings are meaningful and totally interpretable. The only exception appears with the N4 'Self-Consciousness' scale, which loads negatively on the Extraversion factor even though it belongs to the Neuroticism factor. However, this finding is not surprising since according to the theory (Costa and McCrae 1992; Goldberg 1990), self-conscious individuals are usually timid, defensive, shy, and inhibited people opposed to extraverted individuals who are usually bold, assertive, sociable, and energetic individuals. Thus, this negative loading on Extraversion factor, although it is higher than usual, is, to some extent, expected. Additionally, the agreement of the factor solutions derived from the two different samples, as indicated by the congruence coefficients, is very large. More specifically, the congruence coefficients for the

Table 2: Factor loadings for the TPQue sub-scales for the applicant (N = 561) and employee samples (N = 225)

	Applicant Sample					Employee Sample				
TPQue sub-scales	E	N	O	A	C	E	N	O	A	C
Extraversion sub-scales										
Warmth	.46	−.05	.27	.32	.06	.49	.08	.22	.45	−.15
Gregariousness	.62	−.13	.02	.14	.21	.67	−.03	−.14	.10	.07
Assertiveness	.68	−.23	.14	−.02	.38	.71	−.18	.15	−.15	.31
Activity	.69	−.02	.06	.04	.21	.70	−.12	.08	−.02	.14
Excitement Seeking	.74	−.01	.18	.12	.15	.73	−.08	.14	.08	.09
Positive Emotions	.72	−.06	.13	.21	.08	.71	.08	.12	.17	−.08
Neuroticism sub-scales										
Anxiety	−.19	.78	−.03	.01	−.13	−.21	.79	−.07	.10	−.13
Angry Hostility	.07	.71	.11	−.11	−.18	.05	.70	.04	−.16	−.16
Depression	−.34	.72	.02	.14	−.08	−.35	.67	.06	.21	−.32
Self-Consciousness	−.46	.54	−.09	.22	−.06	−.59	.35	−.09	.11	−.20
Impulsiveness	.20	.57	.01	.06	−.20	.17	.57	.05	.10	−.35
Vulnerability	−.31	.73	−.07	.12	−.24	−.26	.72	−.13	.18	−.31
Openness sub-scales										
Fantasy	.09	.27	.36	.16	−.14	−.02	.19	.61	.11	−.16
Aesthetics	.28	.01	.40	.22	.20	−.11	−.05	.51	.46	.07
Feelings	.20	.30	.43	.41	−.05	.06	.28	.49	.41	−.13
Actions	.47	−.10	.41	.06	.21	.34	−.10	.48	−.07	.01
Ideas	.39	−.12	.52	−.07	.13	.31	−.12	.55	−.13	.03
Values	.01	−.02	.59	−.11	−.02	.05	−.12	.51	−.01	−.27
Agreeableness sub-scales										
Trust	.16	.01	−.06	.49	.09	.29	−.08	.03	.44	.01
Straightforwardness	.15	.03	.12	.39	.05	.01	.01	.17	.28	−.03
Altruism	.22	.08	−.02	.57	.12	.05	.11	.01	.56	.01
Compliance	−.18	.03	−.03	.64	.07	−.19	−.11	−.10	.75	−.08
Modesty	−.24	.03	−.04	.38	.18	−.45	−.15	−.07	.38	.01
Tender-Mindedness	.16	.07	.09	.63	.20	.03	.09	−.06	.67	.17
Conscientiousness sub-scales										
Competence	.36	−.29	.06	.09	.60	.21	−.30	−.16	−.09	.72
Order	.03	−.12	−.03	.20	.49	−.08	−.14	−.23	.01	.55
Dutifulness	.20	−.19	.08	.31	.58	.08	−.15	.04	.07	.65
Achievement Striving	.44	−.03	−.09	.11	.66	.37	−.07	.01	−.06	.66
Self-Discipline	.32	−.17	.05	.19	.73	.27	−.17	−.09	.04	.79
Deliberation	.09	−.27	.09	.07	.70	−.18	−.37	−.10	−.01	.69
Factor Congruencies	0.95	0.96	0.97	0.95	0.92					

Note: E = Extraversion, N = Neuroticism, O = Openness, A = Agreeableness, C = Conscientiousness

five factors in both samples are: 0.95 for E, 0.95 for N, 0.94 for O, 0.98 for A, and 0.95 for C.

Furthermore, if we examine the amount of variances explained by each solution, we see that they are almost identical, since in the applicant sample it explains 56,7% of the total variance, while in the non-applicant sample its factor solution accounts for 56,3% of the total variance. The above data support the hypothesis that the five-factor model, as operationalized by the TPQue, remains stable across applicant and non-applicant samples.

In Table 4 the loadings of the TPQue sub-scales on each of the five main factors of the third contrasted pair, employees vs. non-applicants, are presented in addition with the congruence coefficients for the five-factor scales of both samples.

Once more, all the TPQue sub-scales (apart the N4 'Self-Consciousness' scale) were loaded on the intended factor, while the secondary loadings were appropriate and meaningful whenever they appeared. For example, in non-applicant sample, N5 'Impulsiveness' had a large negative secondary loading on Conscientiousness, because according to the theory, people with low self-control are characterized by an inability to manage their impulses or desires (Costa and McCrae 1992). Regarding the amount of variances explained by the two solutions, the similarity was again substantial. In the employee sample the factor solution

Table 3: Factor loadings for the TPQue sub-scales for the applicant (N = 561) and non-applicant samples (N = 1054)

	Applicant Sample					Non-applicant Sample				
TPQue sub-scales	E	N	O	A	C	E	N	O	A	C
Extraversion sub-scales										
Warmth	.46	−.05	.27	.32	.06	.50	−.01	.18	.48	−.02
Gregariousness	.62	−.13	.02	.14	.21	.66	.14	−.10	.15	−.01
Assertiveness	.68	−.23	.14	−.02	.38	.69	.22	.18	−.11	.26
Activity	.69	−.02	.06	.04	.21	.69	.08	.09	−.04	.16
Excitement Seeking	.74	−.01	.18	.12	.15	.71	.11	.16	.05	.11
Positive Emotions	.72	−.06	.13	.21	.08	.73	.12	.08	.12	−.02
Neuroticism sub-scales										
Anxiety	−.19	.78	−.03	.01	−.13	−.13	.83	−.01	.04	−.03
Angry Hostility	.07	.71	.11	−.11	−.18	.20	.64	.08	−.25	−.11
Depression	−.34	.72	.02	.14	−.08	−.33	.72	.04	.09	−.14
Self-Consciousness	−.46	.54	−.09	.22	−.06	−.55	.42	−.16	.15	.12
Impulsiveness	.20	.57	.01	.06	−.20	.27	.36	.03	.02	−.38
Vulnerability	−.31	.73	−.07	.12	−.24	−.27	.73	−.09	.07	−.30
Openness sub-scales										
Fantasy	.09	.27	.36	.16	−.14	.11	.08	.49	.17	−.15
Aesthetics	.28	.01	.40	.22	.20	.13	.01	.55	.22	.14
Feelings	.20	.30	.43	.41	−.05	.19	.21	.45	.43	−.03
Actions	.47	−.10	.41	.06	.21	.40	.09	.43	.06	−.01
Ideas	.39	−.12	.52	−.07	.13	.34	.07	.56	.02	.04
Values	.01	−.02	.59	−.11	−.02	.06	.15	.52	−.08	−.17
Agreeableness sub-scales										
Trust	.16	.01	−.06	.49	.09	.17	.06	−.11	.54	.01
Straightforwardness	.15	.03	.12	.39	.05	.08	.02	.19	.29	.15
Altruism	.22	.08	−.02	.57	.12	.13	.09	.01	.57	.11
Compliance	−.18	.03	−.03	.64	.07	−.32	.08	.06	.64	−.01
Modesty	−.24	.03	−.04	.38	.18	−.32	.08	.05	.38	.11
Tender-Mindedness	.16	.07	.09	.63	.20	.13	.15	.14	.61	.27
Conscientiousness sub-scales										
Competence	.36	−.29	.06	.09	.60	.25	−.31	−.03	.01	.67
Order	.03	−.12	−.03	.20	.49	−.06	−.03	−.16	.07	.50
Dutifulness	.20	−.19	.08	.31	.58	−.01	−.09	.01	.19	.71
Achievement Striving	.44	−.03	−.09	.11	.66	.21	−.01	−.01	−.01	.68
Self-Discipline	.32	−.17	.05	.19	.73	.16	−.12	.03	.05	.79
Deliberation	.09	−.27	.09	.07	.70	−.14	−.20	−.02	.01	.64
Factor Congruencies	0.95	0.95	0.94	0.98	0.95					

Note: E = Extraversion, N = Neuroticism, O = Openness, A = Agreeableness, C = Conscientiousness

accounted for 59,1% while the corresponding amount for the non-applicant sample was 56,3%. Finally, when the congruence coefficients were examined, high resemblance between the two samples was revealed. More specifically, the congruence coefficients for the five factors between the two different samples were: 0.95 for E, 0.96 for N, 0.98 for O, 0.96 for A, and 0.99 for C. These results indicate that there is high similarity between factor patterns in both samples (i.e. employees and non-applicants) as seems to be the case with the other two samples that were contrasted: applicants vs. non-applicants, and applicants vs. employees.

When we ask whether two studies are telling the same story, what we usually mean is whether the results from both studies are reasonably consistent with each other or whether they are significantly heterogeneous. One way to investigate this consistency, when we use correlation coefficients, is to examine whether there is a statistical significant difference between the correlation coefficients derived from the different studies. Table 5 presents the intercorrelations between the main factor scales of the TPQue across the three different samples as well as the estimated *post hoc* power tests between the differences of the correlation coefficients in the three contrasted groups.

For each of the two correlation coefficients compared in each contrasted group, we computed the associated Fisher's z,[1] defined as

Table 4: Factor loadings for the TPQue sub-scales for the employee (N = 225) and non-applicant samples (N = 1054)

	Applicant Sample					Non-applicant Sample				
TPQue sub-scales	E	N	O	A	C	E	N	O	A	C
Extraversion sub-scales										
Warmth	.49	.08	.22	.45	−.15	.50	−.01	.18	.48	−.02
Gregariousness	.67	−.03	−.14	.10	.07	.66	.14	−.10	.15	−.01
Assertiveness	.71	−.18	.15	−.15	.31	.69	.22	.18	−.11	.26
Activity	.70	−.12	.08	−.02	.14	.69	.08	.09	−.04	.16
Excitement Seeking	.73	−.08	.14	.08	.09	.71	.11	.16	.05	.11
Positive Emotions	.71	.08	.12	.17	−.08	.73	.12	.08	.12	−.02
Neuroticism sub-scales										
Anxiety	−.21	.79	−.07	.10	−.13	−.13	.83	−.01	.04	−.03
Angry Hostility	.05	.70	.04	−.16	−.16	.20	.64	.08	−.25	−.11
Depression	−.35	.67	.06	.21	−.32	−.33	.72	.04	.09	−.14
Self-Consciousness	−.59	.35	−.09	.11	−.20	−.55	.42	−.16	.15	.12
Impulsiveness	.17	.57	.05	.10	−.35	.27	.36	.03	.02	−.38
Vulnerability	−.26	.72	−.13	.18	−.31	−.27	.73	−.09	.07	−.30
Openness sub-scales										
Fantasy	−.02	.19	.61	.11	−.16	.11	.08	.49	.17	−.15
Aesthetics	−.11	−.05	.51	.46	.07	.13	.01	.55	.22	.14
Feelings	.06	.28	.49	.41	−.13	.19	.21	.45	.43	−.03
Actions	.34	−.10	.48	−.07	.01	.40	.09	.43	.06	−.01
Ideas	.31	−.12	.55	−.13	.03	.34	.07	.56	.02	.04
Values	.05	−.12	.51	−.01	−.27	.06	.15	.52	−.08	−.17
Agreeableness sub-scales										
Trust	.29	−.08	.03	.44	.01	.17	.06	−.11	.54	.01
Straightforwardness	.01	.01	.17	.28	−.03	.08	.02	.19	.29	.15
Altruism	.05	.11	.01	.56	.01	.13	.09	.01	.57	.11
Compliance	−.19	−.11	−.10	.75	−.08	−.32	.08	.06	.64	−.01
Modesty	−.45	−.15	−.07	.38	.01	−.32	.08	.05	.38	.11
Tender-Mindedness	.03	.09	−.06	.67	.17	.13	.15	.14	.61	.27
Conscientiousness sub-scales										
Competence	.21	−.30	−.16	−.09	.72	.25	−.31	−.03	.01	.67
Order	−.08	−.14	−.23	.01	.55	−.06	−.03	−.16	.07	.50
Dutifulness	.08	−.15	.04	.07	.65	−.01	−.09	.01	.19	.71
Achievement Striving	.37	−.07	.01	−.06	.66	.21	−.01	−.01	−.01	.68
Self-Discipline	.27	−.17	−.09	.04	.79	.16	−.12	.03	.05	.79
Deliberation	−.18	−.37	−.10	−.01	.69	−.14	−.20	−.02	.01	.64
Factor Congruencies	0.95	0.96	0.98	0.96	0.99					

Note: E = Extraversion, N = Neuroticism, O = Openness, A = Agreeableness, C = Conscientiousness

½ log [(1 + *r*)/(1-*r*)]. Next, by applying the Bonferonni *t post hoc* power test, we were able to detect statistically significant differences between the contrasted correlation coefficients. As we can see in Table 5, the majority of the differences in the correlation coefficients were not significant (i.e., invariant) across the various contrasted groups. The minimum difference between correlations (backtransformed from Fisher's *z*) that was detected as statistically significant, was 0.15. Furthermore, all correlation differences are relatively low, suggesting a near-orthogonal relationship among the factors for all the contrasted groups. These results support the stability of the factor structure of the TPQue across the different groups.

Finally, we tested whether the factor solution derived from each contrasted group fits the theoretical framework of the five factor model, as this was operationalized by the TPQue factor structure. Table 6 provides evidence of three goodness-of-fit indices. An absolute index, the chi square/degrees of freedom (χ^2/df), in which figures less than 4.0 (Bollen 1989) indicates acceptable fit. Four relative goodness-of-fit indices, the Tucker-Lewis Index (TLI: Tucker and Lewis 1973), the Normed Fit Index (NFI: Bentler and Bonett 1980), and the Normed Noncentrality Fit Index (CFI: Bentler, 1990), in which values higher than 0.90 indicate a model with a good fit, and the Root Mean Square Error of Approximation (RMSEA: Browne and Cudeck

Table 5: Intercorrelations among the TPQue five factors and post hoc comparisons across the three samples

	E	N	O	A	C
Applicants (N = 561)					
E	–	0	3.25**	2.61	4.75**
N	−.33*	–	−.14	.25	1.75
O	.48*	−.01	–	.75	.25
A	.21*	.08	.18*	–	4.03**
C	.53*	−.41*	.22*	.32*	–
Employees (N = 225)					
E	–	−.27	−2.02	−.78	.27
N	−.33*	–	−.91	.41	−2.43
O	.25*	−.003	–	−.81	−2.16
A	−.002	.06	.12	–	−2.26
C	.21	−.52*	.20*	−.003	–
Non Applicants (N = 1054)					
E	–	−.58	2.11	2.88**	7.69**
N	−.35*	–	.92	.05	−.58
O	.39*	−.07	–	0	5.0**
A	.06*	.03	.18*	–	3.08**
C	.19*	−.38*	−.04	.17	–

Notes: E = Extraversion, N = Neuroticism, O = Openness, **A** = Agreeableness, **C** = Conscientiousness
* $p < .01$, ** $p < .0025$
The intercorrelations between factor scales across the three samples are reported below the diagonal. The post hoc comparisons (Bonferonni *t*) between Fisher's *z* correlation coefficients in applicant vs. employees, employees vs. non-applicants, and applicants vs. non-applicants samples, are reported above the diagonal.

1993), in which values less than 1.0 indicate a model with a good fit. Finally, a parsimonious index, the Parsimonious Normed-Fit Index (PNFI: Mulaik *et al*. 1989) in which values above .80, usually indicate models with good fit.

Results from Table 6 indicate that all three solutions provide a rather acceptable good fit with the theoretical framework of the Five-Factor model (Costa and McCrae 1992). More specifically, the factor structure of the applicant sample fits the data well according to different goodness-of-fit indices (χ^2/df = 2.4, TLI = .97, NFI = .96, CFI = .97, RMSEA = .091, and PNFI = .81). Similarly, the factor structure of the employee sample also fits the data well according to the different goodness-of-fit indices (χ^2/df = 3.3, TLI = .95, NFI = .94, CFI = .95, RMSEA = .099, and PNFI = .80). Finally, the factor structure of the non-applicant sample fits the data well according to the majority of the goodness-of-fit indices (χ^2/df = 2.7, TLI = .95, NFI = .94, CFI = . 96). Both the RMSEA as well as the PNFI goodness-of-fit indices are marginally lower that the minimum requirements, providing at least a suggestive of good fit (RMSEA = .11 and PNFI = .79 respectively). Especially for the RMSEA statistic, this result is not surprising, since it has been argued that it is an index which takes the degrees of freedom into account, penalising

Table 6: Overall fit indices for the TPQue factor scales across samples

Samples	Absolute Indices			Relative Indices				Parsimony Index
	χ^2	Df	χ^2/df	TLI	NFI	CFI	RMSEA	PNFI
Applicants	963.9	391	2.4	.97	.96	.97	.091	.81
Employees	1334.3	395	3.3	.95	.94	.95	.099	.80
Non-applicants	1057.5	390	2.7	.95	.94	.96	.11	.79

Notes: TLI = Tucker-Lewis Index; NFI = Normed Fit Index; CFI = Normed Noncentrality Fit Index; RMSEA = Root Mean Square Error of Approximation; PNFI = Parsimonious Normed-Fit Index.

models that have too many freed parameters (Jackson *et al*. 1996). Since all the models under investigation have many degrees of freedom and most of the parameters are left to run free (instead of specifying the factor loadings for each-sub scale), the RMSEA values were expected to be so close to unity.

Discussion

The results of this study support the stability of the five-factor model of personality across three different settings (applicants, non-applicants, and employees) providing further support to the existence of the five major factors in the personality sphere. Although the results we came up with were in contrast to the findings of a significant study (Schmit and Ryan 1993), they were in a similar vein with the results of Montag and Levin (1994), who found stability of the five factors across two samples of job applicants, in a study carried out in another Mediterranean country, Israel.

The findings also support the construct validity of the personality measure used (TPQue), as a complete measure of the five-factor model in the Greek language. Nevertheless, it should be noted that the use of personality testing in Greece, especially as a selection tool, is still very limited, mainly due to the lack of appropriately developed or adopted instruments (Eleftheriou and Robertson 1999; Kantas *et al*. 1997). For the vast majority of employees and job applicants it was probably the first time they had completed a personality measure, and they were not able to distort their answers in a favourable way, which could produce a factor structure similar to the one identified by Schmit and Ryan (1993). In addition, the lack of a well-established vocational guidance policy (Patiniotis and Stavroulakis 1997), and the scarcity of career counselling centres, which have only recently started to spring up around the country, have kept Greek job seekers away from familiarizing themselves with advanced assessment techniques, such as personality or ability testing.

Summing up, the present study provided further support of the existence of the five personality factors in Greek language, supporting the validity of the TPQue as an adequate measure of Big Five in Greece, across different testing purposes.

Note

1 The advantage of the Fisher's *z* transformation is that equal differences between any pair of Fisher *z*'s are equally detectable, a situation that does not hold for untransformed *r*'s. For raw *r*'s, the difference between .00 and .86, for example (a difference of .86 units of *r* but a difference of 1.3 units of Fisher *z*), is no more detectable than the difference between .86 and .99 (difference of .13 units of *r* but a difference of 1.3 units of Fisher's *z*). In addition, significance tests of difference between *r*'s are more accurate when this transformation is employed (Alexander *et al*. 1989).

References

Alexander, R.A., Sozzaro, M.J. and Borodkin, L.J. (1989) Statistical and empirical examination of the chi-square test for homogeneity of correlations in meta-analysis. *Psychogical Bulletin*, **106**, 329–331.

Barrick, M. and Mount, M. (1991) The big five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, **44**, 1–25.

Barrick, M. and Mount, M. (1993) Autonomy as a moderator of the relationships between the big five personality dimensions and job performance. *Journal of Applied Psychology*, **78**, 111–118.

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

Bentler, P.M. (1990) Comparative fit indexes in structural models. *Psychological Bulletin*, **107**, 238–246.

Bentler, P.M. and Bonett, D.G. (1980) Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, **88**, 588–606.

Berry, J.W. (1969) On cross-cultural comparability. *International Journal of Psychology*, **4**, 119–128.

Bollen, K. (1989) A new incremental fit index for general structural equation models. *Sociometry Methods Research*, **17**, 303–316.

Browne, M.W. and Cudeck, R. (1993) Alternative ways of assessing model fit. In K. A. Bollen. and J. S. Long. (eds) *Testing Structural Equation Modes.* Newbury Park, CA: Sage.

Cellar, D.F., Miller, M.L., Doverspike, D.D. and Klawsky, J.D. (1996) Comparison of factor structures and criterion-related validity coefficients for two measures of personality based on the five fact or model. *Journal of Applied Psychology*, **81**, 694–704.

Costa, Jr, P.T. (1996) Work and personality: use of the NEO-PI-R in industrial/organisational psychology. *Applied Psychology: An International Review*, **45**, 225–241.

Costa, Jr, P.T. and McCrae, R.R. (1989) *The NEO-FFI Manual Supplement*. Odessa, FLA: Psychological Assessment Resources.

Costa Jr, P.T. and McCrae, R.R. (1992) *NEO PI-R. Professional Manual*. Odessa, Florida:

Psychological Assessment Resources, Inc.
Day, D. and Silverman, S. (1989) Personality and job performance: evidence of incremental validity. *Personnel Psychology*, **42**, 25–36.
Digman, J. (1990) Personality structure: Emergence of the five factor model. *Annual Review of Psychology*, **41**, 417–440.
Eleftheriou, A. and Robertson, I. (1999) A survey of management selection practices in Greece. *International Journal of Selection and Assessment*, **7**, 203–208.
Ghiselli, E.E. (1973) The validity of aptitude tests in personnel selection. *Personnel Psychology*, **26**, 461–477.
Goldberg, L.R. (1981) Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (ed), *Review of Personality and Social Psychology* (Vol. 2, pp. 141–165). Beverly Hills, CA: Sage.
Goldberg, L.R. (1990) An alternative 'description of personality': the big-five factor structure. *Journal of Personality and Social Psychology*, **59**, 1216–1229.
Harman, H.H. (1976) Modern Factor Analysis (3rd edn.). London: University of Chicago Press.
Hogan, J. and Hogan, R. (1989) How to measure employee reliability. *Journal of Applied Psychology*, **74**, 273–279.
Hogan, J., Hogan, R. and Busch, C. (1984) How to measure service orientation. *Journal of Applied Psychology*, **69**, 167–173.
Hough, L.M. (1997) The millennium for personality psychology: new horizons or good old daze. *Applied Psychology: An International Review*, **47**, 233–260.
Hough, L.M. and Oswald, F.L (2000) Personnel Selection. *Annual Review of Psychology*, **51**, 631–664.
Jackson, D.N., Paunonen, S.V., Fraboni, M. and Goffin, R.D. (1996) A five factor versus six-factor model of personality structure. *Personality and Individual Differences*, **20**, 33–45.
John, O.P. (1990) The 'Big Five' factor taxonomy: dimension of personality in the natural language and in questionnaires. In L.A. Pervin (ed.), *Handbook of Personality Theory and Research*. New York: Guilford.
Kantas, A., Kalogera, S. and Nikolaou, I. (1997) Recruitment and selection: managerial practices and perceptions in the Greek private sector. In R. Pepermans, A. Buelens, C.J. Vinkenburg, and P.G.W. Jansen (eds), *Managerial Behaviour and Practices: European Research Issues* (pp. 81–90). Leuven: Acco.
Matthews, G. (1997) The big five as a framework for personality assessment. In N. Anderson and P. Herriot (eds), *International Handbook of Selection and Assessment* (pp. 475–492). Chichester: Wiley.
McCrae, R. and John, O. (1992) An introduction to the five-factor model and its applications. *Journal of Personality*, **60**, 175–215.
Montag, I. and Levin, J. (1994) The five-factor personality model in applied settings. *European Journal of Personality*, **8**, 1–11.
Mount, M.K. and Barrick, M.R. (1995) The big five personality dimensions: implications for research and practice in human resources management. In G. R. Ferris (ed.), *Research in Personnel and Human Resources Management* (pp. 153–200). London: JAI Press.
Mount, M.K., Barrick, M.R. and Stewart, G.L. (1998) Five-Factor Model of personality and performance in jobs involving interpersonal interactions. *Human Performance*, **11**, 145–166.
Mulaik, S., James, L., Van Alstine, J., Bennett, N., Lind, S. and Stilwell, C. (1989) Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, **105**, 430–445.
Nikolaou, I. and Robertson, I.T. (in press) The five-factor model of personality and work behaviour in Greece. *The European Journal of Work and Organizational Psychology*.
Ones, D.S. and Viswesvaran, C. (1998) The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance*, **11**, 245–270.
Ones, D.S., Viswesvaran, C. and Reiss, A.D. (1996) Role of social desirability in personality testing for personnel selection: the red herring. *Journal of Applied Psychology*, **81**, 660–679.
Ones, D.S., Viswesvaran, C. and Schmidt, F.L. (1993) Comprehensive meta-analysis of integrity test validities: findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, **78**, 679–703.
Patiniotis, N. and Stavroulakis, D. (1997) The development of vocational education policy in Greece: a critical approach. *Journal of European Industrial Training*, **21**, 192–202.
Robertson, I.T. and Callinan, M. (1998) Personality and work behaviour. *European Journal of Work and Organizational Psychology*, **7**, 317–336.
Rosse, J.G., Stecher, M.D., Miller, J. L. and Levin, R.A. (1998) The impact of response distortion on pre-employment personality testing and hiring decisions. *Journal of Applied Psychology*, **83**, 634–644.
Salgado, J.F. (1997) The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology*, **82**, 30–43.
Salgado, J.F. (1998) Big Five personality dimensions and job performance in Army and civil occupations: a European perspective. *Human Performance*, **11**, 271–288.
Saucier, G., Hampson, S.E. and Goldberg, L.R. (in press). In S. E. Hampson (Ed.), *Advances in personality psychology*, (Vol. 1). London: Routledge.
Schmidt, F.L. and Hunter, J.E. (1998) The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, **124**, 262–274.
Schmit, M. and Ryan, A. (1993) The big five in personnel selection: factor structure in applicant and non-applicant populations. *Journal of Applied Psychology*, **78**, 966–974.
Schneider, R. and Hough, L. (1995) Personality and industrial/organizational psychology. In C. Cooper and I. Robertson (eds), *International Review of Industrial and Organizational Psychology* (Vol. 10, pp. 75–129). Chichester: Wiley.
Tett, R., Jackson, D. and Rothstein, M. (1991) Personality measures as predictors of job performance: a meta-analytic review. *Personnel*

Psychology, **44**, 703–742.
Tsaousis, I. (1996) The psychometric assessment of personality in Greek speaking population. Unpublished doctoral dissertation, University of London Goldsmith's College, London, U.K.
Tsaousis, I. (1999) The traits personality questionnaire (TPQue): A Greek measure for the five factor model. *Personality and Individual Differences*, **26**, 271–283.
Tucker, L. and Lewis, C. (1973) A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, **38**, 1–10.