



Searching for the “Big Five” in a Greek context: the NEO-FFI under the microscope

Georgia Panayiotou ^{a,*}, Constantinos M. Kokkinos ^b, George Spanoudis ^a

^a *University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus*

^b *Democritus University of Thrace, Nea Chili, 68100 Alexandroupoli, Greece*

Received 21 January 2003; received in revised form 7 June 2003; accepted 25 July 2003

Abstract

The five-factor inventory (FFI) is a commonly used personality test based on the NEO-personality inventory revised. It has been translated into several languages and validated in a number of countries. Using a sample of 1204 individuals, the present study evaluates the psychometric properties and factor structure of the Greek FFI and provides normative information for its use with Greek populations. Convergent validity is also assessed by studying the relationship between the five scales of the FFI and the Brief Symptom Inventory. Results show that although the factor scales had acceptable internal consistency, they were highly intercorrelated. Exploratory factor analyses failed to reproduce the appropriate factor structure, yielding instead numerous fragments of the five dimensions. Confirmatory factor analyses also failed to lend support to the five factor model of personality as measured by this instrument. Based on these results, the use of the instrument is recommended only with serious caution in this cultural context. Further research ought to decipher whether the failures to reproduce the five factor model using the FFI in this and other cultures represents a challenge to the universality of the theory, or merely a shortcoming of the specific instrument.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: NEO-FFI; Assessment; Personality; Big Five

1. Introduction

One of the most popular conceptions of personality today, which rests on the assumption of a specific number of underlying personality dimensions, is the five factor model. A significant body

* Corresponding author. Tel.: +357-22-753267; fax: +357-22-377950.

E-mail address: georgiap@ucy.ac.cy (G. Panayiotou).

of research supports the presence of five orthogonal dimensions of personality, which apparently also exist in the trait adjectives people use in their daily lives to describe others or themselves (Goldberg, 1993). The “Big Five” dimensions that have emerged from factor analyses of peer and self-ratings of personality descriptors are neuroticism, extraversion, openness to experience, conscientiousness and agreeableness. They have been replicated in several countries including Korea, France, Estonia and Finland (e.g. Pulver, Allik, Pulkkinen, & Hamalainen, 1995; Rolland, Parker, & Stumpf, 1998; Spirrison & Choi, 1998) using various personality inventories (Piedmont, McCrae, & Costa, 1991).

The most popular assessment instrument for these five dimensions is Costa and McCrae’s NEO-personality inventory revised (NEO-PI-R), which consists of 240 items that result in the five factors and a number of factor facets. Because of its length, a briefer version of the instrument, the NEO-five factor inventory (NEO-FFI), was created (Costa & McCrae, 1992a) with 60 items. To derive each of the factor scales of the NEO-FFI, 12 items from the 1986 administration of the NEO-PI, were selected for their high positive loading on the corresponding trait. The brief instrument has adequate internal reliability and correspondence with the full scale (McCrae & Costa, 1989) and thus has been used widely for research and clinical purposes. It shows good convergent validity with other personality instruments, correlates with ratings provided by others of the target’s personality (Kurtz & Sherker, 2003; Parker & Stumpf, 1998), and appears to be a valid measure of genetically/biologically based personality dimensions (Riemann, Angleitner, & Strelau, 1997).

In spite of its widespread use and the increasing acceptance of its external validity, research carried out so far in a number of countries using exploratory and confirmatory factor analyses (CFA), has called into question the factor structure of the NEO-FFI with some studies finding fewer than five factors (Ackerman & Heggstad, 1997; Ferguson & Patterson, 1998), others finding as many as eight (Yoshimura, Ono, Nakamura, Nathan, & Suzuki, 2001) and yet others finding weak fit indices for a five factor solution (Mooradian & Nezlak, 1996). This is not surprising given that even the long version of the instrument has not fared well when subjected to CFA (e.g. Parker, Bagby, & Summerfeldt, 1993).

Concern regarding the use of the FFI has surfaced both in English and non-English speaking countries. In a British study, Egan, Deary, and Austin (2000) found that although British norms corresponded favourably with American, the factor structure of the NEO-FFI was less satisfactory. Factors N, A and C appeared reliably, while many of the items of factors O and E did not load adequately on the expected factors. In a Canadian sample of female college students, Holden and Fekken (1994) found acceptable support for the five factors with exploratory factor analysis, but failed to obtain a five factor solution with CFA. Even with an American sample, Tokar, Fischer, Snell, and Harik-Williams (1999) failed to obtain satisfactory orthogonal or oblique factor solutions with CFA. Hence, the utility of the NEO-FFI in measuring the “Big Five” dimensions needs to be evaluated cautiously even with regards to English speaking cultures where no translation of the instrument was required.

The need to empirically validate this instrument is even more pressing in non-English speaking nations where the questionnaire requires translation. In a study carried out in Germany, Schmitz, Hartkamp, Baldini, Rollnik, and Tress (2001) examined the psychometric properties of the German NEO-FFI among a group of psychosomatic patients and a reference group of personality disorder patients. Their findings are of concern to FFI users: First, the five dimensions were not orthogonal (the same concern was raised by Church & Burke (1994)). Secondly, CFA failed to

yield acceptable fit indices for five, four or two factor models using both oblique and orthogonal rotations. Similarly, Yoshimura et al. (2001) obtained eight factors with a Japanese sample, after subjecting the NEO-FFI to exploratory factor analysis, with 11 items not loading on any of the factors.

The five-factor model of personality has yet to receive adequate empirical attention in the Greek cultural context, partly due to the absence of well-documented measures of the construct. Hence, the present study examines the internal reliability and factor structure of the Greek translation of the NEO-FFI. This translation has been initially authored by Bezevengis for use by parents to rate their children's personality (see Stalikas, Triliva, & Roussi, 2002) at the University of Athens, and was further developed by Demetriou and Kazi (2001) at the University of Cyprus for use with adults, but to date no research to validate the instrument or establish Greek norms has been published. The method of front and back translation was used for its development and the measure is used in research in both Greece and Cyprus (Demetriou, Kyriakides, & Avraamidou, in press). The present authors proceeded to an independent back translation of the instrument, which resulted in a very close approximation to the original that was deemed satisfactory.

The present study additionally generates some initial norms based on a large sample of participants. Finally, the study provides evidence regarding the concurrent validity of the instrument in relation to the Brief Symptom Inventory (BSI; Derogatis, 1993). Both the NEO-FFI and the BSI are frequently used in mental health settings. Being able to demonstrate that the NEO-FFI is able to predict the presence of psychological symptoms (i.e. deviations from 'normal' personality functioning) in our sample, would add credence to its validity. For instance, one should expect that extraversion would be associated with low scores on anxiety, depression, somatization, etc., whereas neuroticism would be associated with high scores on these BSI dimensions.

2. Method

2.1. Participants

Participants were a total of 1204 individuals who had taken part in four previous studies conducted in Cyprus that included the NEO-FFI as one of their measures. These studies were as follows: Study 1 was on primary school teachers' professional burnout and included 447 individuals. Study 2 was on stress among secondary school teachers in Cyprus, and included 345 individuals. Study 3 was on student teachers' attitudes toward pupils' undesirable behaviors. This study included 48 students from Cyprus and 94 students from Greece. Finally, study 4 examined marital and premarital relationships among Cypriots and included 135 heterosexual couples, i.e. 270 individuals. With regards to education, the samples from the two first studies were of a similar educational level (all had at least an undergraduate degree) whereas the sample from studies 3 and 4 was educationally heterogeneous. To verify that this sample heterogeneity does not reflect different underlying personality factor structures, analyses were conducted to explore any personality differences between more and less educated participants (see Section 3.5). Demographic characteristics of the full sample were as follows: 375 were male and 827 were female (two participants did not report their sex). Participants' age ranged from 15–78, with a mean of 33.72 (SD = 12.6). Place of residence (rural or urban) was allowed to vary randomly in all samples. On

the basis of these characteristics the samples were deemed to be adequately similar so as to permit subsequent analyses to be performed for the sample as a whole.

An additional sample of 123 individuals took the NEO-FFI and the BSI in order to obtain convergent validity information. This sample (79 females, 44 males) was of a mixed educational level and ranged in age from 16–70 ($M = 28$, $SD = 12$).

2.2. Measures

The NEO-FFI (Costa & McCrae, 1992a) is an instrument that contains 60 NEO-PI-R questions measuring five dimensions of the normal personality, i.e. neuroticism, extraversion, openness to experience, agreeableness and conscientiousness. This abbreviated instrument is somewhat limited compared to the long version in that it does not provide facet information. Respondents indicate their degree of agreement with each item on a five point Likert type scale (0 = strongly disagree, 4 = strongly agree). The Greek instrument is a direct translation of the NEO-FFI maintaining the same number of items and scoring procedures as the original. Scale scores were obtained by summing the 12 items of each dimension, after appropriate items were reversed (Costa & McCrae, 1992a).

Participants in the convergent validity sample additionally took the BSI (Derogatis, 1993). This is a 53-item self-report symptom inventory, which is a brief version of the SCL-90-R (Derogatis & Cleary, 1977) a well-studied instrument that is frequently used in clinical and research settings. It covers nine main symptom dimensions, namely somatization, obsessive compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation and psychoticism. Each item is scored on a five point Likert type scale (0 = not at all, 4 = extremely) on which the respondent indicates the degree to which each symptom has caused him/her distress during the last seven days.

2.3. Procedure

All participants responded to the 60 items of the Greek NEO-FFI. Means, standard deviations and Cronbach's alpha reliability coefficients were calculated for each of the five subscales of the instrument (N, E, O, A, C). Furthermore, means and standard deviations by sex were calculated. The NEO-FFI items were then subjected to factor analytic procedures, including varimax and oblique rotations as well as CFA, to verify the factor structure of the instrument. For all exploratory factor analyses, a loading of 0.30 was used as a cut-off point. Convergent validity was assessed by evaluating the bivariate correlations between the NEO-FFI dimensions and the BSI subscales for the 123 participants.

3. Results

3.1. Psychometric properties

Table 1 shows the means, standard deviations (also broken down by sex) and Cronbach's alpha reliability coefficients for each of the five scales of the NEO-FFI and compares them with the results reported by Costa and McCrae (1992a).

Table 1
Psychometric properties and internal reliability of the Greek NEO-FFI

	N	E	O	A	C
Mean (all subjects)	33.16	40.38	37.65	44.62	47.36
SD (all subjects)	7.12	6.30	5.56	5.98	7.10
Alpha	0.78	0.70	0.51	0.69	0.83
Alpha (US norms)	0.86	0.77	0.73	0.68	0.81
Mean (males/present sample)	30.80	40.96	36.76	43.91	47.49
SD (males/present sample)	6.70	6.20	5.34	6.22	7.19
Mean (males/US norms)	17.60	27.22	27.09	31.93	34.10
Mean (females/present sample)	34.19	40.13	38.05	44.94	47.31
SD (females/present sample)	7.06	6.32	5.60	5.84	7.01
Mean (females/US norms)	20.54	28.16	26.98	33.76	35.04
Mean (ages ≤ 22/present sample)	35.44	39.73	37.30	41.75	43.37
SD (ages ≤ 22/present sample)	8.01	6.63	5.85	6.21	7.26
Mean (ages >22/present sample)	32.51	40.57	37.75	45.43	48.50
SD (ages >22/present sample)	6.73	6.19	5.49	5.66	6.57

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

Separate means and standard deviations are also reported for college aged young adults (ages ≤ 22) and older adults (ages >22), as in the US sample significant differences were observed between these populations. In the studies reported in the Costa and McCrae (1992a) manual, college students tend to be somewhat higher than adults on N, E and O and somewhat lower on A and C. Only the latter was observed in our sample, with N, E and O being very similar across the two age groups. Means for all scales, across the whole sample, were somewhat higher than those reported by Egan et al. (2000) and by Schmitz et al. (2001). Reliability coefficients were satisfactory except for the Openness scale, which was 0.51.

A further inspection of the distributions of scores of each dimension revealed that agreeableness and conscientiousness were negatively skewed, whereas neuroticism was positively skewed, indicating perhaps an attempt by the respondents to present themselves in a favorable light. To exclude the possibility that this effect was due to the high education level of a significant number of participants, skewness statistics were recalculated after dividing the sample into two categories: The primary and secondary school teachers (from studies one and two respectively) were grouped together because they all had at least a college education and the same profession. The student teachers in study three and the mixed education group in study 4 were also placed together, based on the assumption that this was a more randomly selected population in terms of their education. They included a large section of less than college-educated individuals (about 80% of this sample had either a high school or some college education but no college degree). Skewness statistics were similar in both groups, except for the Openness dimension (see Table 2), indicating that the attempt to present oneself in a favourable light was common across all participants.

3.2. Scale independence

Pearson correlations were calculated for all scale pairs, in order to examine whether the scales are orthogonal. Results, as presented in Table 3, indicate a high degree of association between

Table 2
Skewness statistics for participants grouped according to education level

FFI dimensions	N	E	O	A	C
College-educated participants ($N = 792$)	0.34	-0.08	-0.13	-0.55	-0.45
Participants with less than college education ($N = 412$)	0.37	-0.21	0.14	-0.31	-0.35

Table 3
Pearson correlations among the subscales of the Greek NEO-FFI

	Neuroticism	Extraversion	Openness	Agreeableness
Extraversion	-0.45**			
Openness	-0.05	0.23**		
Agreeableness	-0.28**	0.20**	0.12**	
Conscientiousness	-0.25**	0.30**	0.15**	0.33**

Note: **Correlation is significant at the 0.01 level.

several of the scales: N was negatively correlated with E, A and C. E was positively correlated with O, A and C and finally O was positively correlated with A and C.

When scale scores were used as items, exploratory factor analysis with varimax rotation resulted in a single factor that explained 39.4% of the variance, with positive loadings for all dimensions except N, with a negative loading of -0.69. This was done in an attempt to replicate the finding of Egan et al. (2000) who, using a similar analysis, found that the five factors were reduced to a two-factor solution, where one dimension represented a non-psychopathological personality type and a second represented “something other than general adjustment.” Our findings did not replicate these two primary types of adjustment, and indicate that the instrument evoked a one-dimensional super-ordinate construct of personality: one is either well adjusted (open, agreeable, sociable, conscientious) or not (neurotic).

3.3. Group differences

One-way ANOVAs were used to examine subscale mean differences between men and women. Women scored significantly higher than men on neuroticism, $F(1, 1148) = 57.56$, $p < 0.001$, openness, $F(1, 1132) = 13.27$, $p < 0.001$, and agreeableness, $F(1, 1154) = 7.48$, $p < 0.006$. Men on the other hand scored higher on extraversion, $F(1, 1152) = 4.36$, $p < 0.04$. No significant gender differences were found on the conscientiousness subscale.

In order to examine whether the education level of the participants affected their scores on the five dimensions, additional one-way ANOVAs were carried out comparing those with and without a college education. Group 1 (primary and secondary school teachers) scored significantly lower on neuroticism, $F(1, 1150) = 22.42$, $p < 0.001$, and significantly higher on openness, agreeableness and conscientiousness, $F(1, 1134) = 12.50$, $p < 0.001$, $F(1, 1156) = 138.69$, $p < 0.001$, and $F(1, 1133) = 97.14$, $p < 0.001$ compared to Group 2 (mixed education). There were no significant differences between the two groups in extraversion.

3.4. Norms

Given the large size of the sample, normative values based on *T*-score conversions of the means for each scale can be calculated using the formula $T = 10 \times [(\text{Raw score} - \text{Mean Raw Score} / \text{SD of Raw Scores})] + 50$. These can provide a general guideline for the use of the instrument in the Greek cultural context, although users should be cautious given the large representation in the sample of college graduates, and specifically teachers.

3.5. Principal components analysis

To examine the factor structure of the Greek NEO-FFI, firstly a principal components analysis with varimax rotation (assuming orthogonality of the subscales) was conducted. Initially, separate analyses were conducted to confirm sample homogeneity: college-educated and less than college-educated individuals had similar factor structures, as indicated by the scree plots for each sample that resulted in five primary factors (similar scree plots were obtained for teachers and non-teachers and for Cypriot and Greek individuals). On this basis, it was deemed appropriate to treat the sample as a whole for all the analyses that follow.

Fourteen factors with eigenvalues over 1 were extracted, which explained 54.4% of the variance. Scree plot inspection indicated the presence of five primary factors that explained 35.8% of the variance. Inspection of the content of these five factors indicated little relationship to the expected five dimensions. However, since the scree plot pointed to the presence of five factors, and given that the internal reliability coefficients of the five a priori dimensions were relatively high, a five-factor solution was further explored. The factor analysis was re-run specifying a five-factor model. The extracted five factors accounted for 36% of the variance having eigenvalues over 2. Factor 1 was mostly comprised of conscientiousness, factor 2 of neuroticism, factor 3 of extraversion, factor 4 of agreeableness and factor 5 of openness items. However, only 40 of the 60 items belonged to their appropriate factors and 18 items had cross-loadings over 0.30 on inappropriate factors (see Table 4).

Because the scales are not entirely independent, the NEO-FFI items were further subjected to an oblique rotation. The extracted factors again bore little resemblance to the anticipated five dimensions, though scree plot inspection indicated five primary factors. Examination of the emergent factors indicated that the anticipated five may have become fragmented into subsections: There were two fragments of C (5 and 5 items each) three fragments of N (5, 2 and 5 items), two fragments of O (5 and 2 items), two fragments of A (4 and 6 items) and two fragments of E (4 and 4 items each—see Table 5). In addition to these 11 factors, three single item factors also emerged. However, it must be noted that the fragments were not “pure” in terms of content, i.e. they also contained items that theoretically should belong to other factors. Thus, 10 items loaded on inappropriate factors.

These results led to the question as to whether the anticipated five dimensions had been broken down along theoretically meaningful “facets”, despite the fact that the NEO-FFI does not lend itself to facet analysis. This hypothesis was rejected on the basis of two findings: First, visual comparison of the obtained fragments with the original facets contained in the NEO-PI-R (Costa & McCrae, 1992a) showed little resemblance. Additionally, a second-order factor analysis with oblique rotation was carried out on the 11 obtained subdivisions to see if the fragments would

Table 4

Factor loadings and cross-loadings for specified five-factor solution: Exploratory factor analysis using a varimax rotation

Item number	Factor to which item belongs	Observed factor loadings				
		C	N	E	A	O
50	C	0.71				
35	C	0.71				
25	C	0.66				
20	C	0.66				
40	C	0.63				
60	C	0.62				
5	C	0.58				
10	C	0.47				
55	C	0.47	0.36			
45	C	0.44	0.40			
49	A	0.43		0.34		
4	A	0.43			0.36	
8	O	-0.42				
15	C	0.39	0.32			
47	E	0.38				
53	O	0.35		0.33		0.32
38	O					
26	N		-0.67			
51	N		-0.63			
11	N		-0.63			
21	N		-0.60			
42	E		0.60			
41	N		-0.59			
6	N		-0.56			
12	E		0.52			
56	N		-0.52			
36	N		-0.50		-0.39	
30	C	0.34	0.43			
18	O					
17	E			0.64		
37	E			0.59		
2	E			0.56		
7	E			0.56		
22	E			0.49		
32	E	0.31		0.49		
52	E	0.43		0.43		
34	A	0.32		0.42		
46	N			-0.40		
28	O			0.38		
16	N			-0.31		0.31
14	A				0.65	
44	A				0.60	
54	A				0.54	
24	A				0.47	
59	A				0.42	
29	A				0.41	

Table 4 (continued)

Item number	Factor to which item belongs	Observed factor loadings				
		C	N	E	A	O
58	O			0.35	-0.37	
57	E				0.36	
19	A	0.32			0.35	
9	A	0.30	0.33		0.33	
27	E				0.33	
48	O					0.56
33	O					0.52
43	O			0.34		0.47
23	O					0.46
31	N					0.44
13	O					0.42
1	N					0.41
39	A					0.41
3	O					0.32

Note: Only factor loadings >0.30 are shown.

Table 5

Factor fragments emerging from oblique rotation of the 60-item Greek FFI

Factor fragments	Item number
C1	25, 5, 60, 35, 20
C2	15, 30, 45, 55, 10
N1	11, 21, 1, 41, 31
N2	16, 46
N3	6, 26, 56, 51, 36
O1	43, 13, 53, 23, 58
O2	33, 48
A1	19, 34, 49, 4
A2	44, 24, 29, 14, 54, 59
E1	2, 17, 22, 7
E2	52, 37, 32, 47

Note: Table shows only those items that loaded on their theoretically anticipated factor fragments.

load on their appropriate five factors. This resulted in four meaningless factors, which failed to reflect the anticipated dimensions.

3.6. Confirmatory factor analysis

The raw data were also subjected to CFA using the M-plus 2.02 program (Muthén & Muthén, 2001). Based upon the entire sample, maximum likelihood estimation method was applied. A model was chosen in which the five factors were orthogonal and each item loaded only on its corresponding factor. Chi-square statistic was 17,865.25 based on 1700 degrees of freedom, which was not statistically significant, as would be expected in a large sample. Several indices of fit were calculated all of which suggested that the five factor model misfits the data: CFI was 0.56, well below 0.90, which is typically considered evidence of good fit (Mueller, 1996). Conversely,

RMSEA was 0.064, SRMR was 0.095, which are higher than the cut-points recommended by Hu and Bentler (1999).

Due to the fact that the exploratory factor analysis using a specified five factor solution had been found to explain 36% of the variance with 40 items, an additional five factor model was tested using CFA with only these 40 items. Again, indices of fit were unacceptably low, verifying that the Greek FFI does not yield a five factor solution (CFI = 0.71; RMSEA = 0.064; SRMR = 0.083). Because, however, in the exploratory factor analyses a number of items loaded primarily on factors other than the ones theoretically anticipated, and loaded secondarily on their anticipated factor, the CFA was re-run allowing for those cross-loadings (see Table 4). Indices of fit were also not indicative of a five factor model, CFI = 0.62, RMSEA = 0.062, SRMR = 0.09.

To further explore a five factor solution, we followed the recommendations of McCrae, Zonderman, Bond, and Paunonen (1996), who conducted a series of CFA using facet scores instead of item scores, allowing for facets to inter-correlate. They assumed that although the five factors are theoretically orthogonal, the facets in reality contain items that are related to each other. However, one must keep in mind that the NEO-FFI does not permit facet analysis since items were not selected so that all facets are represented in the instrument. Also, McCrae et al. (1996) used cross-loadings of facets based on their normative sample data. No such normative information exists for the Greek NEO-FFI. Due to these limitations we calculated facet scores for only those facets that had at least one item represented. We obtained the correlations between these facets and the five dimensions through exploratory factor analysis (varimax rotation). First, we entered facet scores into a CFA assuming a simple five-factor model, where each facet loads only on its appropriate factor. This resulted in an improvement of fit indices, although these remained unacceptable, CFI = 0.69, RMSEA = 0.085, SRMR = 0.09. Given that it is feasible to make allowances for secondary loadings by fixing the loadings at values other than zero, we fixed all cross-loadings using the correlations obtained from the exploratory analysis, in a complete model. This resulted in CFI = 0.88, RMSEA = 0.053, and SRMR = 0.056. Hence, although using facets' secondary cross-loadings improves fit indices somewhat, the five-factor model continues not to be supported by our data.

3.7. Convergent validity

Table 6 shows the correlations between the five subscales of the NEO-FFI and the nine scales of the BSI. As is evident, neuroticism was positively and significantly associated with all symptom

Table 6
Correlations among personality and symptom subscales

	SOM	O-C	ISD	DEP	ANX	HOS	PHOB	PAR	PSY
N	0.42**	0.51**	0.56**	0.68**	0.57**	0.38**	0.48**	0.50**	0.56**
E	-0.05	-0.14	-0.25**	-0.35**	-0.18	-0.13	-0.22*	-0.19*	-0.28**
O	0.02	-0.02	-0.02	0.12	-0.09	-0.04	-0.06	0.01	-0.07
A	-0.25**	-0.21*	-0.34**	-0.37**	-0.31**	-0.57**	-0.20*	-0.52**	-0.43**
C	-0.14	-0.44**	-0.26**	-0.27**	-0.17	-0.17	-0.28**	-0.03	-0.35**

Note: SOM = Somatization; O-C = Obsessive compulsive; ISD = Interpersonal sensitivity; DEP = Depression; ANX = Anxiety; HOS = Hostility; PHOB = Phobic anxiety; PAR = Paranoid ideation; PSY = Psychoticism.

*Significant at the $p < 0.05$ level.

**Significant at the $p < 0.01$ level.

scales, whereas the other four personality dimensions were either not associated or negatively associated with the symptom scales. This finding verifies that neuroticism measures a “negative” or possibly pathological aspect of personality and that the other scales measure positive aspects of personality as they increase in value. However, these results also reinforce the finding reported earlier that the five personality scales in this culture actually converge to a single super-ordinate dimension of personality that ranges from “well-adjusted” to “maladjusted.”

4. Discussion

The short form of the NEO-PI-R, the 60-item FFI, has been used extensively, but its psychometric properties continue to be scrutinized. Several studies have cast doubt on its ability to measure the same five dimensions measured by its longer predecessor (Ackerman & Heggestad, 1997; Ferguson & Patterson, 1998; Yoshimura et al., 2001). The present study examines the psychometric properties and factor structure of the Greek NEO-FFI, provides normative information and tests its convergent validity vis a vis the BSI. Scale internal consistency was acceptable when the items were a priori allocated to their corresponding five factors. However further analyses raised serious concerns rendering the utility of the instrument questionable. These results add to the cautiousness that has been expressed in studies conducted in both English and non-English speaking countries.

More specifically, two primary concerns were raised having to do with (a) the orthogonality of the scales and (b) the factor structure of the instrument. The correlations conducted among the scales, when a priori accepted, showed a high degree of inter-relations. Though these are somewhat meaningful, with N negatively related to the “positive” personality traits (i.e. C, A, O and E) and the latter positively related to each other, they do not reflect the theorized independence of the scales (Costa & McCrae, 1992b; Goldberg, 1990). A second level factor analysis showed that the five scales loaded on a single factor, supporting further their inter-correlation. Several authors have raised questions regarding the orthogonality of the scales measuring the Big Five (e.g. Becker, 1999). As Saucier (2002) puts it “there is a discrepancy between theory and practice” p.2, with many of the scales measuring the Big Five, such as the NEO-FFI resulting in significant correlations among the dimensions in many studies. Hence, the present study is not unique in coming across this problem, but it adds to the general concern about the adequacy of scales such as the NEO-FFI to appropriately measure the five factor model, and casts doubt on the ability of the Greek version of the instrument to measure these dimensions in this cultural context.

The second and main issue of concern has to do with the factor structure of the instrument. Neither the varimax nor the oblique rotation produced satisfactory five factor solutions. Both analyses yielded 14 factors, which explained a fairly small percentage of the variance and left out a significant proportion of the items (at least 13) not loading on the appropriate factors. Even when a five factor solution was specified, 20 items did not load on the expected dimensions, and many items loaded highly on more than one factors. The results of a second-order factor analysis did not support that the 14 factors were fragments of the original dimensions. Furthermore, when subjected to CFA, a five factor model failed to reach acceptable levels of fit. Bearing in mind that the 60 items of the NEO-FFI were selected on the basis of their high loadings on their

corresponding factors, the lack of support for a five-factor solution in this study constitutes reason for serious concern regarding the validity of this instrument.

The findings of the present study also bring into question the issue of the cultural validity of the Greek FFI. The a priori dimensions seemed to converge on a single super-ordinate factor that appears to correspond to the notion of adjustment: That is, one can be either neurotic or well adjusted. It seems implausible that within the Greek culture this is the only meaningful aspect of personality. Instead, it is more likely that the Greek FFI is a poor measure of the way people are described and describe themselves and of the five factor model specifically. Perhaps, the only viable approach to producing the five dimensions, if they are indeed as universal as it is claimed by supporters of this model, is for one to return to the original lexical approach, and find out which adjectives people in particular cultures use to identify meaningful personality traits.

The concerns raised in this study have ramifications that go beyond the issue of cultural appropriateness. Ideally, one would expect the FFI to be as good a measure of the “Big Five” as its predecessor, the NEO-PI-R, which has fared well in a variety of cultures. This has not proven to be the case. The present study adds to a significant body of research carried out in both English and non-English speaking countries (Egan et al., 2000; Holden & Fekken, 1994) that questions the ability of the instrument to confirm the existence of the Big Five dimensions. Hence, it appears that the instrument begs for revision. Unfortunately, in our sample the problem of cross-loadings and items failing to load on their appropriate factors was so extensive that it is questionable whether any part of this instrument would be worth salvaging.

It appears, then, that simple translation and/or adaptation of the NEO-FFI in different populations is an inadequate approach to creating a reliable and valid measure of the “Big Five” personality dimensions. A better strategy would perhaps be to translate and use the NEO-PI-R and, assuming that the five factors are reliably obtained, choose the best representatives of each scale to create culturally appropriate short instruments. Alternatively, the NEO-PI-R can be used as a foundation, enriched with novel, culturally appropriate items. From this, short measures can then be obtained.

In the meantime, the use of the Greek FFI can only be recommended with serious caution. Because of the large sample size, the norms suggested in this study are believed to be stable, though probably more appropriate for individuals with at least some college education. Our sample was admittedly opportunistic, and hence characteristics such as socio-economic status, education and age may have influenced score distributions. Further studies could extend these results to more diverse populations, if the instrument is to continue to be used in its current form. Perhaps the most significant lesson to be learned from this study is that simple translation of a personality instrument, no matter how widely validated this has been, is an inadequate method for deriving a culturally valid measure of underlying personality structures.

References

- Ackerman, P., & Heggestad, E. (1997). Intelligence, personality and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*, 219–245.
- Becker, P. (1999). Beyond the Big Five. *Personality and Individual Differences*, *26*, 511–530.

- Church, A., & Burke, P. (1994). Exploratory and confirmatory tests of the Big Five and Tellegen's three—and four-dimensional models. *Journal of Personality and Social Psychology*, *66*, 93–114.
- Costa, P. T., & McCrae, R. R. (1992a). *The NEO PI-R professional manual*. Odessa, Florida: Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1992b). Reply to Eysenck. *Personality and Individual Differences*, *13*, 667–673.
- Demetriou, A., & Kazi, S. (2001). *Unity and modularity in the mind and the self: Studies on the relationships between self-awareness, personality, and intellectual development from childhood to adolescence*. London: Routledge.
- Demetriou, A., Kyriakides, L., & Avraamidou, C. (in press). The missing link in the relations between intelligence and personality. *Journal of Research in Personality*.
- Derogatis, L. R. (1993). *BSI brief symptom inventory. Administration, scoring and procedures manual*. Minneapolis: National Computer Systems.
- Derogatis, L. R., & Cleary, P. A. (1977). Confirmation of the dimensional structure of the SCL-90-R: A study in construct validation. *Journal of Clinical Psychology*, *33*, 981–989.
- Egan, V., Deary, I., & Austin, E. (2000). The NEO-FFI: emerging British norms and an item-level analysis suggests N, A, and C are more reliable than O and E. *Personality and Individual Differences*, *29*, 907–920.
- Ferguson, E., & Patterson, F. (1998). The five factor model of personality: Openness a distinct but related construct. *Personality and Individual Differences*, *12*, 773–790.
- Goldberg, L. R. (1990). An alternative “description of personality”: The Big-Five factor structure. *Journal of Personality and Social Psychology*, *59*, 1216–1229.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*, 26–34.
- Holden, R. R., & Fekken, G. C. (1994). The NEO Five-Factor Inventory in a Canadian context: Psychometric properties for a sample of university women. *Personality and Individual Differences*, *17*, 441–444.
- Hu, L. T., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.
- Kurtz, J. E., & Sherker, J. L. (2003). Relationship quality, trait similarity, and self-other agreement on personality ratings in college roommates. *Journal of Personality*, *71*, 21–48.
- McCrae, R. R., & Costa, P. T. (1989). Rotation to maximize the construct validity of factors in the NEO Personality Inventory. *Multivariate Behavioral Research*, *24*, 107–124.
- McCrae, R. R., Zonderman, A. B., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the revised NEO Personality Inventory: Confirmatory factor analysis versus proctustes rotation. *Journal of Personality and Social Psychology*, *70*, 552–566.
- Mooradian, T. A., & Nezlak, J. B. (1996). Comparing the NEO-FFI and Saucier's Mini-Markers as measures of the Big Five. *Personality and Individual Differences*, *21*, 213–215.
- Mueller, R. O. (1996). *Basic principles of structural equation modelling*. New York: Springer Verlag.
- Muthén, L. K., & Muthén, B. O. (2001). *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.
- Parker, J. D. A., Bagby, R. M., & Summerfeldt, L. J. (1993). Confirmatory factor analysis of the revised NEO personality inventory. *Personality and Individual Difference*, *15*, 463–466.
- Parker, W. D., & Stumpf, H. (1998). A validation of the five-factor model of personality in academically talented youth across observers and instruments. *Personality and Individual Differences*, *25*, 1005–1025.
- Piedmont, R. L., McCrae, R. R., & Costa, P. T., Jr. (1991). Adjective check list scales and the five factor model. *Journal of Personality and Social Psychology*, *60*, 630–637.
- Pulver, A., Allik, J., Pulkkinen, L., & Hamalainen, M. (1995). A Big Five personality inventory in two non-Indo-European languages. *European Journal of Personality*, *9*, 109–124.
- Riemann, R., Angleitner, A., & Strelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self- and peer-report NEO-FFI scales. *Journal of Personality*, *3*, 449–475.
- Rolland, J. P., Parker, W. D., & Stumpf, H. (1998). A psychometric examination of the French translations of the NEO-PI-R and NEO- FFI. *Journal of Personality Assessment*, *7*, 269–291.
- Saucier, G. (2002). Orthogonal markers for orthogonal factors: The case of the Big Five. *Journal of Research in Personality*, *36*, 1–31.
- Schmitz, N., Hartkamp, N., Baldini, C., Rollnik, J., & Tress, W. (2001). Psychometric properties of the German version of the NEO-FFI in psychosomatic outpatients. *Personality and Individual Differences*, *31*, 713–722.

- Stalikas, A., Triliva, S., & Roussi, P. (2002). *Ta psychometrica ergalia stin Ellada [Psychometric tools in Greece]*. Athens: Ellinika Grammata.
- Spirrison, C. L., & Choi, S. (1998). Psychometric properties of a Korean version of the revised NEO-Personality Inventory. *Psychological Reports, 83*, 263–274.
- Tokar, D. M., Fischer, A. R., Snell, A. F., & Harik-Williams, N. (1999). Efficient assessment of the five-factor model of personality: Structural validity analyses of the NEO five-factor inventory (Form S). *Measurement and Evaluation in Counselling and Development, 32*, 14–30.
- Yoshimura, K., Ono, Y., Nakamura, K., Nathan, J. H., & Suzuki, K. (2001). Validation of the Japanese version of the NEO-five factor inventory in a large community sample. *Psychological Reports, 88*, 443–449.