

# Psychometric properties of the International Personality Item Pool Big-Five personality questionnaire for the Greek population

Maria Ypofanti,<sup>1,2</sup> Vasiliki Zisi,<sup>2</sup>  
Nikolaos Zourbanos,<sup>2</sup>  
Barbara Mouchtouri,<sup>3</sup>  
Pothiti Tzanne,<sup>1</sup> Yannis Theodorakis,<sup>2</sup>  
Georgios Lyrakos<sup>1,4</sup>

<sup>1</sup>Department of Psychiatry, General Hospital Nikaia-Piraeus Agios Panteleimon, Athens; <sup>2</sup>Department of Physical Education and Sport Science University of Thessaly; <sup>3</sup>Psychiatric Hospital (Daphni), Athens; <sup>4</sup>Second Department of Anesthesiology, School of Medicine, University of Athens, Pain Unit, Attikon University Hospital, Athens, Greece

## Abstract

Goldberg's International Personality Item Pool (IPIP) big-five personality factor markers currently lack validating evidence. The structure of the 50-item IPIP was examined in two different adult samples (total N=811), in each case justifying a 5-factor solution, with only minor discrepancies. Age differences were comparable to previous findings using other inventories. One sample (N=193) also completed additionally another personality measure (the TIPI Short Form). Conscientiousness, extraversion and emotional stability/ neuroticism scales of the IPIP were highly correlated with those of the TIPI ( $r=0.62$  to  $0.65$ ,  $P=0.01$ ). Agreeableness and Intellect/Openness scales correlated less strongly ( $r=0.54$  and  $0.58$  respectively,  $P=0.01$ ). The IPIP scales have good internal consistency ( $\alpha=0.88$ ) and relate strongly to major dimensions of personality assessed by the two questionnaires.

## Introduction

Personality is considered a system of distinctive characteristics and developmentally dynamic procedures that influence psychological functioning of every individual. It has even been suggested as a mediator of the environmental stressors on the onset and development of an illness.<sup>1,2</sup>

The assessment of personality has been proved an important issue in the field of psychology research. A widely used and accepted

as a reliable questionnaire for the assessment of personality characteristics was the Revised NEO Personality Instrument (NEO PI-R),<sup>3</sup> which was based on the notion that there are only five basic dimensions of personality.<sup>4</sup> Although the NEO PI-R has been used as a tool in numerous studies,<sup>5-8</sup> it was argued that the five factors personality model was an arbitrary selection procedure among interactions of experiential factors.<sup>9</sup> The model was actually aiming to incorporate concepts related to an individual's predisposition,<sup>7,10-12</sup> and some researchers, supporters of the theory of the model of five factors, argued that personality can be perceived through five general factors of predisposition,<sup>13-15</sup> a view that was soon reflected in a new assessment instrument.

The NEO PI-R was replaced by the International Personality Item Pool (IPIP) questionnaire,<sup>15</sup> which has two versions: a full version with 100 items and a short version with 50 items. The later version, which is assessed in the present study, includes five factors: i) extroversion-introversion, ii) agreeableness, iii) conscientiousness, iv) emotional stability-neuroticism, v) intellect. Each factor has 10 items/questions, which receive a positive or a negative score. Mlacic and Goldberg weighed the questionnaire for the Croatian population,<sup>16</sup> using both the full and short versions and found an almost identical to the American structure. Further investigation indicated that the reliability of the short version was from acceptable to excellent, making this short edition, a practical tool at the area of personality assessment.<sup>17</sup>

The issue of the type and number of personality factors is still under investigation. Some researchers supported that personality can be described with the use of fewer than five factors,<sup>18</sup> while others doubted that five factors are enough to perceive the total perspective of the personality function.<sup>19,20</sup> It was also supported the need of about seven factors (like dependency and honesty or the preparedness to take risk), in order to achieve a reliable personality assessment.<sup>21,22</sup> The whole portrait of an individual's personality however, has applications in many areas, and although the personality five factor model remains open to investigation as a tool of assessment,<sup>23</sup> the IPIP has been used in several studies. For example, it has been used in measuring religious gratitude and wellbeing,<sup>24-26</sup> in examining the effects of personality traits on team processes and outcomes,<sup>27</sup> on non-medical use of prescription drugs,<sup>28</sup> even in exploring how an investor's dispositional affect and cognitive style influence venture investment portfolio concentration.<sup>29</sup> Finally, we should mention that public managers are aware of personality assessment, use it in their jobs, and are generally convinced of its efficacy.<sup>30</sup>

Since the IPIP's inception, portions of the

Correspondence: Georgios N. Lyrakos, General Hospital Nikaia, Piraeus "Agios Panteleimon", Department of Psychiatry, D. Mandouvalou 3, Nikaia, 18454, Greece.  
Tel.: +30.2132.077646 - Fax: +30.2132.077143.  
E-mail: geolr@hotmail.com

Key words: Personality; IPIP Big-Five factor markers; validation.

Acknowledgments: we would like to thank all the translators that helped in the validation procedure as well as all the individuals who agreed to participate in the study.

Contributions: the authors contributed equally.

Conflict of interest: the authors declare no potential conflict of interest.

Received for publication: 24 December 2013.

Revision received: 7 March 2014.

Accepted for publication: 7 March 2014.

This work is licensed under a Creative Commons Attribution NonCommercial 3.0 License (CC BY-NC 3.0).

©Copyright M. Ypofanti et al., 2015  
Licensee PAGEPress, Italy  
Health Psychology Research 2015; 3:2206  
doi:10.4081/hpr.2015.2206

item pool have been translated, or are in the process of being translated, into Arabic, Bulgarian, Chinese, Croatian, Danish, Dutch, Estonian, Finnish, French, German, Hebrew, Hmong, Hungarian, Italian, Korean, Latvian, Norwegian, Persian, Polish, Romanian, Russian, Serbian, Slovene, Spanish, Swedish, Turkish, Vietnamese, and Welsh. A page at the IPIP Web site keeps the research community informed about such translation projects (including multiple projects in Chinese, German, Spanish, and Swedish), with e-mail links to the investigators involved.<sup>15</sup> No Greek translation however, is listed on this web page.

The purpose of the present study was to translate and measure the psychometric properties of this questionnaire in order to obtain a useful and valid instrument for use in the public services of Greece, because: i) it is cost free, ii) its items can be obtained instantaneously via the Internet, iii) it includes over 2000 items, all easily available for inspection, iv) scoring keys for IPIP scales are provided, and v) its items can be presented in any order, interspersed with other items, reworded, translated into other languages, and administered on the World Wide Web without any permission request.

## Materials and Methods

### Participants

The sample size was calculated to be representative of the Greek adult population, which is estimated to be 11 million, assuming a confidence level of 99% and an accuracy level that would only tolerate a difference of 0.05. The sample size was estimated at 664 individuals, but assuming a 10% level of missing data, we aimed to recruit 850 patients.

To help obtain a heterogeneous sample, a snowball recruitment procedure was used; 10 colleagues, five from two hospitals in Athens and five from two University Departments in Thraki (Komotini) and Thessaly (Trikala) served. Collected data from persons they knew as well as from persons these acquaintances knew. Each colleague tried to find 5 volunteers who would then ask two relatives, friends, or colleagues to complete the questionnaire. From the 1000 estimated sample only 850 adults returned the questionnaires. Since only participants who were born and had lived continuously in Greece and who were fluent in Greek were invited to take part in the study, the final sample consisted of 811 participants of Greek ethnicity. The sample consisted of 281 men and 530 women with a mean age of 35 (SD=11.2). The age range in the sample was from 17 to 75 years. The majority of these participants (60.4%) had college degrees as their highest educational qualification, with smaller groups having attained secondary schooling (29.1%), primary schooling (5.1%) or post-graduate (5.4%) degrees. In terms of marital status, 49% were married, (44.8%) were single, (5.4%) were divorced and the remainders were widows (0.7%).

### International Personality Item Pool

The IPIP Big-Five factor markers consist of a 50-item and 100-item inventory which can be freely downloaded from the internet.<sup>15</sup> The current study makes use of the 50-item version consisting of 10 items for each of the Big-Five personality factors: Extraversion (E), Agreeableness (A), Conscientiousness (C), Emotional Stability (ES), and Intellect (I). We administered the IPIP items with a 5-point, Likert-type scale ranging from 1 (very inaccurate) to 5 (very accurate) as in the original instrument.<sup>1</sup> Ten additional items were added in the Greek version from the IPIP item pool for discriminate validity reasons and in order to be able to select items, identify dimensions, and measure reliability and internal and concurrent validity as it is proposed in the IPIP site ([www.ipip.ori.org](http://www.ipip.ori.org)).

The internal consistency of the factors (Cronbach  $\alpha$ ) reported by Goldberg,<sup>15</sup> are between 0.79 to 0.87. Below are examples of questions that evaluate each factor, such as

extraversion (E-1: Am the life of the party), agreeableness (A-2: Feel little concern for others), conscientiousness (C-43: Follow a schedule), emotional stability - neuroticism (ES-4: Get stressed out easily) and intellect (I-5: Have a rich vocabulary). The answers to the 60 questions were given in a 5 point Likert scale, from (1) strongly disagree to (5) agree.

### Ten Items Personality Intex, 2006

In order to investigate the cross-sectional validity of the Greek version of the IPIP another short questionnaire measuring the same five factors of personality was also administered to the population of the study. This was the Ten Items Personality Intex-(TIPI),<sup>31</sup> which was also translated into Greek following the same 2 way forward and backward translation and was validated and used in a previous study with diabetic patients.<sup>32</sup> TIPI assesses the same aforementioned five major factors of personality with two questions for each factor respectively. The TIPI uses a 7-point Likert-type scale ranging from 1 (disagree strongly) to 7 (agree strongly). In the present study Cronbach  $\alpha$  was 0.46 for emotional stability, 0.55 for extraversion, 0.52 for openness to experience, 0.39 for agreeableness and 0.52 for conscientiousness, results almost the same with the German version of the questionnaire.<sup>33</sup> The convergent validity for the TIPI with the 44-item BFI was comparable to the convergent validity of other longer multi-item FFM measures.<sup>34</sup> Discriminant correlations were substantially lower than convergent correlations and the TIPI demonstrated good stability as indexed by test-retest correlations.<sup>35</sup>

### Translation of the instrument

The translation strategy was based on minimal criteria developed by the Scientific Advisory Committee of the Medical Outcomes Trust.<sup>35</sup> Translation was performed using the multiple forward and backward translation protocol recommended by Guillemin.<sup>36</sup> Following these, two independent bilingual health professionals translated the questionnaire into Greek (forward translation).

The mother tongue of all translators was the Greek language and their level of English was advanced. A reconciliation meeting was conducted to obtain a consensus version. In this meeting the two translations were compared and found differences only in item p46 (I am quiet around strangers), which was not similar in the two translations (in the first translation it was silent while in the second quiet). Finally the translators decided to keep the quiet version for face validity reasons. It was also decided that one of the reversed items, (p22: Am not interested in other people's problems) would be used in a positive direction in order to be more easily understood.

Then, two native English (living and work-

ing in Greece for the last 10 years) who were blinded to the original version retranslated the re-conciliated Greek version into the source language (back translation) which is the recommended procedure for creating semantic equivalence.<sup>37</sup>

The last step of the translation procedure was the pre-testing of the translated instrument in a small population of hospital staff following a cognitive debriefing process as it is explained by Lyrakos *et al.*<sup>38</sup> This process refers to an in-depth interview of the individuals about their understanding of the questionnaire with the purpose of revealing inappropriate items and translation alternatives. Namely, after completing the questionnaire participants gave their general impression on the clarity of the items, the relevance of the content to their situation, the comprehensiveness of the instructions and their ability to complete it on their own. The same issues were addressed to them for every single item and they were able to make suggestions whenever necessary.

### Item selection

A first selection of items was made from the descriptive response distribution for each item. The criteria used to guide item selection/deletion were as follows: high rates of nonresponse and *not applicable* response ( $\geq 20\%$ ), except for items where high rates in this response category were expected, ceiling, and floor effects ( $\geq 50\%$ ), and unacceptable item total correlation for each factor independently with values less than 0.200.

### Data analyses

The normality of the items of all measures was investigated and found to be within the level recommended for confirmatory factor analysis CFA with maximum-likelihood (ML) estimation (skewness<sup>2</sup>, kurtosis<sup>7</sup>) and still within acceptable values for normality.<sup>39,40</sup>

The psychometric properties of the Greek version of the IPIP were analyzed as follows: i) item-level exploratory factor analysis (EFA) was implemented in order to evaluate the proposed five-factor structure in the Greek adaptation of the IPIP items. Principal component analysis (PCA) with an orthogonal (Varimax) rotation was utilized to assess the internal structure of the measure. Confirmatory factor analysis (CFA) via a structural equation modeling approach,<sup>41</sup> using maximum likelihood methodology was then conducted in the sample in order to evaluate the two models (5 factors each but with different item loadings and number of items). Goodness of fit was assessed following Dragioti *et al.*'s methodology for CFA<sup>42</sup> using both measures of absolute and relative fit. The chi square of each model is reported, but due to its sensitivity to the sample size,<sup>44</sup> the relative chi square ( $\chi^2/df$ )<sup>43</sup> is also provided. Three further absolute meas-

ures of fit are reported, namely: goodness-of fit index (GFI),<sup>35</sup> Bollen's relative fit index (RFI)<sup>41</sup> and root mean square error of approximation (RMSEA).<sup>45</sup>

The fit of each model compared to the null one (*i.e.*, the model that assumes the co-variation among the indicators is due to chance) was assessed by using two relative fit indices, namely the non-normed fit index (NNFI) and the comparative fit index (CFI);<sup>46,47</sup> ii) internal consistency reliability of the instrument was assessed using Cronbach's  $\alpha$  coefficient (1951),<sup>48</sup> and corrected item-total correlations; and iii) convergent validity was assessed by examining the relationship using a Pearson's product-moment correlation between the IPIP and the ten item personality scale (TIPI).

To evaluate the principal component analysis (PCA), clusters of items were observed and interpreted. Criteria for retaining extracted component(s) included: i) Eigenvalue of one or greater;<sup>49</sup> ii) Percentage of variance accounted for by the retained component(s); and iii) Screeplot.<sup>50</sup> Item-component correlations of 0.4 and above were retained. Alpha coefficients of 0.70 or higher and corrected item-total correlations higher than 0.40 were deemed to indicate good reliability.<sup>51,52</sup>

## Results

### Descriptive statistics

Out of the 850 adults who were selected, 811 participated (95.4%): 281 (34.6%) males and 530 (65.4%) females. Their age ranged from 18 to 72 (mean 37.4, SD 11.29), they had 0-4 children (Mean $\pm$ SD: 0.89 $\pm$ 0.98), found it difficult to answer in 1-4 items of the IPIP (Mean 1.4), understood 56-59 questions of the 60 IPIP questions (Mean 57.6) (Table 1).

Of them, 41 (5.1%) had attended primary school, 236 (29.1) high school, 490(60.4%) had an undergraduate degree, 36 (4.4%) a Master of Science (M.Sc.) and 8 (1%) a Ph.D. Regarding marital status, 363 (44.83%) were singles, 398 (49%) married, 44 (5.42%) divorced and 6 (0.75%) widowed. The place of residence was the capital city of Athens in 439 (54.1%) and providence in 372 (45.9%).

The mean and standard deviations of the IPIP and the subscales of each dimension of personality aroused from the factor analysis, as well as for the TIPI and the retest evaluation of the IPIP are shown in Table 1 respectively.

### Scale reliability

#### Item selection

The first omission of items, due to low item total correlation for the whole 60 item questionnaire, led to the abstract of five items, item C-3(H2590) ( $r=0.093$ ), A-7(H21) ( $r=0.086$ ), ES-9(E141) ( $r=0.002$ ), and E-36(X68)

( $r=0.100$ ). Finally item E-46(H661) was found to have a negative loading even though it was supposed to be positive and was decided to be omitted from the questionnaire ( $r=-0.146$ ).

Then, Cronbach's  $\alpha$ , were also calculated for each factor independently in order to omit the additional two items from each factor since a fifty item solution is proposed from the constructors of the original questionnaire, which led us to the omission of five more items with item total correlation less than 0.200 for each factor were also excluded from the questionnaire. These items were, I-10(X176) ( $r=0.178$ ), C-13(H1362) ( $r=0.133$ ), ES-19(X156) ( $r=0.183$ ), I-20(H1230) ( $r=0.203$ ) and A-47 (H107) ( $r=0.186$ ). This led to the final 50 item questionnaire that was further investigated for this study (all omitted items as well as the final Greek translation are presented in Supplementary Tables S1 and S2).

#### Internal consistency

Internal reliability coefficient for the total score of the IPIP questionnaire was 0.882 (0.866 for the 60 item questionnaire) which showed that the scale has very good internal consistency. The internal consistency coefficient (Cronbach's  $\alpha$ ) for each factor of the Greek version of the IPIP was: for Conscientiousness 0.875, for Emotional Stability/Neuroticism 0.849, for Intellect 0.780, for Agreeableness 0.758 and for Extraversion 0.791.

Item total correlations ranged between 0.21 (in agreeableness) and 0.73 (in conscientiousness), respectively. Cronbach's  $\alpha$ , for each factor independently, both for 60 and 50 items, as well as the item total correlations are shown in Table 2 respectively.

#### Test-retest reliability

Out of 80 post graduate students randomly selected as a convenience subsample of the original sample to facilitate evaluation of external (test-retest) reliability, eventually 68 participated (15% dropout). Cronbach's  $\alpha$ , for this subset at the initial test period was 0.88, and was reduced by 0.05 at the retest period ( $\alpha=0.83$ ). Correlations between the test and retest mean scores of these participants were likewise from moderate to high ( $r=0.36-0.67$ ,  $P<0.001$ ; Table 3), suggesting that test-retest reliability was good for the Greek version of the IPIP.

#### Construct validity: scale and subscales intercorrelations

In order to evaluate construct validity of the IPIP scale all the internal correlations were calculated and were found to be above 0.3. Only the Intellect factor had a higher correlation with the Agreeableness  $r=0.322$  ( $P<0.001$ ) and Extraversion factors of the IPIP,  $r=0.56$  ( $P<0.01$ ). None of the other factors correlated higher than this (Table 3).

### Internal structure: exploratory factor analysis

The assessment of the sampling adequacy diagnostics resulted in satisfactory measures of Sampling Adequacy values (0.43 to 0.76). Furthermore, Bartlett's test of sphericity ( $\chi^2=9377.3$ ,  $df=1225$ ,  $P<0.001$ ) indicated the intercorrelations among the items were satisfactory. The Kaiser-Meyer-Olkin index was 0.86, indicating low partial intercorrelations among items. The scree plot produced suggested the extraction of 6 factors accounting for 43.7% of the variance with the sixth factor adding less than 5% in the total variance. This is not reported here (details are available from the authors on request).

As the 6-factor solution results mainly from a split of the agreeableness factor into two and the sixth factor variance was unsatisfactory, in order to examine whether the item loadings were in accordance with the test construction and theory, 5 factors accounting for 42.6% of the variance were extracted from the a random sample of 500 out of the 811 of the IPIP data (the loadings shown in Table 4) by PCA and subjected to varimax rotation. Five eigenvalues above 1.5 were found (5.1, 4.8, 4.4, 3.6, and 3.3) each one explaining at least an additional 5% in the total 42.6% variance (10.3%, 9.7%, 8.9%, 7.1% and 6.7% respectively according to Schönrock-Adema criteria for factor analysis).<sup>44</sup>

Two additional models with seven and four factors were also examined due to the previous criteria were found to have factor complexity higher than one in most cases and both the models explained lower total variance than the five factor model.<sup>44</sup> Salient loadings ( $>0.20$ ) on factors different than the proposed ones were present, while in some cases the loading of an item was very similar in magnitude on a different factor than the proposed one. Further investigation of the questionnaire's latent structure was implemented via confirmatory factor analysis as follows. Based on these results the data can be considered suitable for factor analysis.

All 10 conscientiousness items loaded over 0.40 on the same factor. The same happened in the Emotional Stability items and with the Intellect items. In the conscientiousness items, with 1 item had lower cross-loading of  $>0.30$  in the Intellect items. Eight of the Agreeableness items loaded on the same factor, whilst one item, A-42(E136) loaded almost the same with the Intellect items, and one item, A-57 (H1) loaded in the Extraversion factor. Eight of the Extraversion items had their highest loading on the same factor with one lower cross-loading in E-56-(H592); Finally one item, E-26(H909) loaded highest with the Emotional stability items, with a smaller, but  $>0.40$  loading in the same factor, while E-51(H1110) had its highest loading with the intellect items.

### Confirmatory factor analysis

In accordance with other research concerning five-factor inventories a model with correlated factors as well as a simple structure model was specified.<sup>15-17</sup> The relative and

absolute fit measures indicated that the five-factor model fit the data inadequately but was more acceptable than the independent and the model with no intercorrelations among the latent variables (Table 4). Chi square was sta-

tistically significant ( $\chi^2=3223.20$ ,  $P<0.001$ ), while the relative chi square ( $\chi^2/df$ ) was 2.76, not exceeding very much the rule-of-thumb threshold (close to 2) for acceptable fit.<sup>43</sup> The values of RFI (0.63) and RMSEA (0.058) also

**Table 1. Means and standard deviation of the questionnaires.**

	N.	Min	Max	Mean	SD	Skewness		Kurtosis	
						Mean	SE	Mean	SE
Age	520	18	77	37.41	11.29	0.571	0.107	0.020	0.214
Conscientiousness	519	11	64	38.71	8.98	-0.684	0.107	-0.174	0.214
Emotional stability	519	10	50	29.45	8.44	-0.022	0.107	-0.534	0.214
Intellect	519	17	50	38.62	6.15	-0.445	0.107	0.341	0.214
Agreeableness	518	21	52	42.82	5.50	-0.936	0.107	0.680	0.214
Extraversion	519	10	59	35.31	7.60	-0.456	0.107	0.060	0.214
Conscientiousness retest	68	21	50	40.56	7.44	-0.794	0.291	-0.044	0.574
Emotional stability retest	68	10	45	29.68	7.27	-0.398	0.291	0.366	0.574
Intellect retest	68	29	50	39.96	4.58	0.032	0.291	-0.332	0.574
Agreeableness retest	68	31	50	43.54	4.63	-0.712	0.291	0.086	0.574
Extraversion retest	68	17	50	37.28	7.05	-0.233	0.291	0.088	0.574
Conscientiousness TIPI	194	1	7	5.72	1.25	-1.060	0.175	0.893	0.347
Emotional stability TIPI	194	1	7	4.54	1.47	-0.237	0.175	-0.544	0.347
Intellect TIPI	194	1	7	5.26	1.28	-0.626	0.175	-0.100	0.347
Agreeableness TIPI	194	2	7	5.60	1.18	-0.746	0.175	0.078	0.347
Extraversion TIPI	194	1	7	4.37	1.43	-0.198	0.175	-0.541	0.347

TIPI, Ten Item Personality Inventory; Min, Minimum; Max, Maximum; SD, Standard Deviation; SE, Standard error.

**Table 2. Cronbach's  $\alpha$  and item total correlations for the final International Personality Item Pool factors.**

	Cronbach $\alpha$ 60 items	Cronbach $\alpha$ 50 items	Range of item total correlations, r	Omitted items
Conscientiousness	0.856	0.875	0.403-0.725	C3-H250, C13-H1362
Emotional stability	0.836	0.849	0.416-0.692	ES9-E141, ES19-X156
Intellect	0.762	0.780	0.353-0.519	I10-X176, I20-H1230
Agreeableness	0.702	0.758	0.210-0.573	A7-H21, A47-H107
Extraversion	0.754	0.791	0.359-0.555	E36-X68, E46-H661
Total	0.866	0.882		

**Table 3. Product moment Pearson correlations between IPIP and TIPI.**

	Conscientiousness	Emotional stability	Intellect	Agreeableness	Extraversion	Age
Conscientiousness	1	0.140**	0.191**	0.248**	0.066	0.168**
Emotional stability	0.140**	1	0.188**	0.090*	0.278**	-0.025
Intellect	0.191**	0.188**	1	0.322**	0.567**	-0.044
Agreeableness	0.248**	0.090*	0.322**	1	0.280**	0.015
Extraversion	0.066	0.278**	0.567**	0.280**	1	-0.106*
Conscientiousness*	0.616**	0.202	0.233	0.068	-0.005	0.141
Emotional stability*	0.047	0.678**	0.099	0.103	0.067	0.138
Intellect*	0.094	0.153	0.501**	0.015	0.402**	-0.062
Agreeableness*	0.126	0.175	0.006	0.452**	0.237	-0.199
Extraversion*	0.046	-0.102	0.119	0.123	0.355**	-0.128
Extraversion**	-0.201**	0.044	0.429**	0.095	0.618**	-0.045
Agreeableness**	0.218**	0.137	0.061	0.540**	0.003	0.231**
Conscientiousness**	0.654**	0.177*	0.196**	0.236**	-0.032	0.094
Emotional stability**	0.288**	0.637**	0.165*	0.208**	0.148*	0.067
Intellect**	0.020	0.122	0.584**	0.204**	0.411**	-0.084
Age	0.168**	-0.025	-0.044	0.015	-0.106*	1

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

suggested an adequate fit, according to Browne and Cudeck (1993).<sup>45</sup> The values of the relative fit measures were moderate (NNFI=0.67, CFI=0.75); values higher than 0.90 indicate close fit. Since the above results suggested the initial model did not fit the data adequately in  $\chi^2$ , alternative models were assessed. Similar indices indicating inadequate fit were found. In order to identify a model with adequate fit, a series of CFA models was implemented by omitting items from the initial scale via a sequential procedure. The item exclusion criteria were: i) low communalities or high loadings to more than one factor (factor complexity higher than one); ii) low item-total correlations (removing items that reduced the Cronbach's  $\alpha$  coefficient); and finally, iii) CFA fit indices of the resulting models. This procedure concluded to a shorter (47 items) version that provided less satisfactory fit indices as can be seen in Table 4. All possible one-factor models and a global factor model also were evaluated again for the 10 items of each factor independently; in all cases the fit was unsatisfactory and thus these results are not reported here. This led us to accept the 50 items model with five factors correlating with one another as it was the most acceptable of all (Supplementary Figure S1).

### Convergent validity

Correlations between IPIP factors and the TIPI factors are presented in Supplementary Table S3, based on 250 participants. There were clear one-to-one relations between all five corresponding factors in both the questionnaires. The IPIP-TIPI Extraversion correla-

tion is  $r=0.62$  ( $P<0.01$ ). Concerning the Emotional Stability/Neuroticism scale scores, the IPIP-TIPI, correlated  $r=0.64$  ( $P<0.01$ ). The IPIP-TIPI Agreeableness correlation is  $r=0.54$  ( $P<0.01$ ). The Conscientiousness scale scores from the IPIP and the TIPI correlated  $r=0.65$  ( $P<0.01$ ). The 5th factors of the IPIP and TIPI (Intellect and Openness respectively) correlated  $r=0.58$  ( $P<0.01$ ). Associations were examined within each of the measures.

### Sex, age, education and location comparisons

Independent sample t test was calculated in order to reveal significant differences between males and females. Analysis shown that there was a significant difference in three out of the five factors of the IPIP questionnaire, Conscientiousness, Emotional Stability and Agreeableness ( $t=-2.856$ ,  $t=3.175$  and  $t=-6.166$   $P<0.05$ ) respectively (Table 5).

One way analysis of variance (ANOVA) with Bonferonni correction was applied to the sample in order to explore possible differences in the factors of IPIP for education and marital status. Significant differences were found in the Conscientiousness factor between single ( $M=36.7\pm 9.4$ ) and married people ( $M=40.5\pm 8.1$ ) Mean Difference (MD)=-3.795,  $P<0.001$ . In education, significant differences were found in Extraversion (MD=-4.6,  $P=0.028$ ) between people having 9 years of compulsory education ( $M=30.9\pm 8.5$ ) and people with postgraduate degrees ( $M=35.6\pm 7.5$ ). Also, significant differences were found in Emotional Stability between people with 9

years of compulsory education ( $M=24.2\pm 9.2$ ) and people with 12 years of education ( $M=29.2\pm 8.4$ ) (MD=-5.01,  $P=0.49$ ), as well as with undergraduates ( $M=29.8\pm 8.2$ ) (MD=-5.63,  $P=0.01$ ) and Master graduates ( $M=32.7\pm 7.9$ ) (MD=-8.50,  $P=0.004$ ) respectively. Finally according to residency, independent sample t test shown a significant difference between residents of Athens ( $M=34.9\pm 7.8$ ) and providence ( $M=37.8\pm 6.2$ ) ( $t=-3.766$ ,  $P<0.001$ ) in Extraversion.

Regarding the age of the sample, Pearson r correlations were also calculated with IPIP and TIPI. There was a significant but low correlation in the Conscientiousness factor with age ( $r=0.168$ ,  $P<0.001$ ) and a significant but negative correlation between age and Extraversion ( $r=-0.106$ ,  $P<0.05$ ) while in TIPI there was a significant correlation between Agreeableness and age ( $r=0.231$ ,  $P<0.001$ ).

## Discussion and Conclusions

This study attempted to validate the IPIP Big-Five markers in Greece. The results of the current study provided substantial support for the generalizability of the 5-factor IPIP structure in a Greek context.

Our results confirmed the factor structure proposed by Goldberg,<sup>1</sup> both in a big sample, for the short IPIP scale. Only minor deviations from the expected item loadings occurred in EFA and CFA analysis but this was expected since the same has happened in all the translations of the IPIP in other countries as well.

**Table 4. Confirmatory factor analysis for all the EFA models that were used in exploratory analysis.**

Models	$\chi^2$	Df	$\chi^2/df$	RFI	RMSEA (P-close)*	NNFI	CFI
50 items with intercorrelations	3223.20	1165	2.767	0.636	0.058 (0.000)	0.667	0.755
Independent model	9684.40	1275	7.596	0.000	0.113 (0.000)	0.000	0.000
50 items with no intercorrelations	3560.88	1175	3.031	0.601	0.063 (0.000)	0.632	0.716
47 items with all intercorrelations	3531.63	1024	3.449	0.687	0.055 (0.000)	0.715	0.778
47 items (with p42 in F3)	3551.83	1024	3.469	0.685	0.055 (0.000)	0.714	0.776
47 items with no intercorrelations	3937.63	1034	3.808	0.654	0.059 (0.000)	0.683	0.734

\*P under the null hypothesis is that RMSEA equals 0.05, the threshold for close fit.

**Table 5. Independent sample t-test for sex.**

	Males (N=281)		Females (N=530)		t	P-value
	Mean	SD	Mean	SD		
Conscientiousness	37.17	9.39	39.56	8.63	-2.86	<i>0.005</i>
Emotional stability	31.02	8.51	28.58	8.29	3.18	<i>0.002</i>
Intellect	38.9.9	6.07	38.42	6.20	1.02	0.311
Agreeableness	40.8.2	5.76	43.94	5.03	-6.17	<i>0.001</i>
Extraversion	35.9.1	7.31	34.99	7.75	1.33	0.185

SD, Standard Deviation; t, test; P value, significant values in italics.

For example, in the Chinese version, only 48 items were loading in the expected factors,<sup>54</sup> while in the Scottish version, 6 of the Agreeableness items loaded together, with a further 3 loading highest with the Extraversion items.<sup>55</sup>

Regarding CFA, even though our results were not perfect, the 50 item model show comparatively better fit than its alternatives. At the item level, fit for all models was less adequate (e.g., for the 50 item model the CFI was 0.76, the NNFI was 0.67, and the SRMSE was 0.58). This was not particularly surprising, given previous research findings regarding the difficulty of conducting CFA using item level data.<sup>56,57</sup>

The reliabilities of the IPIP scales were high except for the factor of Agreeableness but even there results did not exceed Kline's criterion of 0.70. The 50 items correlated highly with the 10 items of the TIPI and the relations between the two questionnaires revealed explicitly clear one-to-one relations between all five corresponding factors in the 50 items version. All these results suggested that the IPIP Big-Five factor markers have an accepted structural validation. Similar results have been reported in other studies as well, like in the Chinese version of the questionnaire,<sup>53</sup> where a similar low Cronbach  $\alpha$ , was found in the Agreeableness factor as well. In contrast with our results, in the Croatian validation of the IPIP the lowest Cronbach  $\alpha$ , was found in the conscientiousness factor.<sup>16</sup>

Although, the results of the current study supported the 5-factor IPIP structure in the Greek sample, they were not perfect. Specifically, the Agreeableness factor might be improved in a next research with a different sample and additional items that will be tested in a new CFA procedure. In our research we had only Caucasian participants since one criteria for participating in the study was the Greek language

In our sample, women were found to have significant differences from men in the Conscientiousness as well as in the agreeableness and in the emotional stability factor, result that complies with the findings of other researchers like the Scottish,<sup>54</sup> as well as Ehrhart *et al.*,<sup>57</sup> who found support for the invariance of the factor structure across groups, although he found some evidence of differences across gender and ethnic groups for model parameters. Despite these differences, since we observed that the proportion of women in the finally recruited sample is twice the size of men the observed differences among age might be ought to this discrepancy and the ratio of highest educational status might prominently elevate also. This was one of the limitations of this study since a big sample of men did not return the questionnaires thus the expected equal variance among the sample for males and females was not

achieved.

In conclusion even though there are limitations to the study, as mentioned above, the results show that construct validity, internal consistency, and concurrent validity of the Greek version of the IPIP, and its corresponding subscales, were generally supported by our population; thus, the 50-item IPIP seems to be a valid tool assessing personality in general population in Greece.

## References

1. Goldberg LR. A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In: Mervielde I, Deary I, De Fruyt F, Ostendorf F, eds. *Personality psychology in Europe*, vol. 7. Tilburg: Tilburg University Press; 1999. pp. 7-28.
2. Van Heck GL. Personality and physical health: toward an ecological approach to health-related personality research. *Eur J Pers* 1997;11:415-43.
3. Costa PT Jr, McCrae RR. Revised neo personality inventory and neo five-factor inventory professional manual. Odessa: Psychological Assessment Resources; 1992.
4. Wiggins JS, Pincus AL. Personality: structure and assessment. *Annu Rev Psychol* 1992;43:473-504.
5. Paunonen SV, Jackson DN, Trzebinski J, Forsteling F. Personality structure across cultures: a multimethod evaluation. *J Pers Soc Psychol* 1992;62:447-56.
6. Benet-Martinez V, John OP. Los cinco grandes across cultures and ethnic groups: Multitrait multimethod analyses of the big five in Spanish and English. *J Pers Soc Psychol* 1998;75:729-50.
7. Jang KL, McCrae RR, Angleitner A, et al. Heritability of facet-level traits in a cross-cultural twin sample: Support for a hierarchical model of personality. *J Pers Soc Psychol* 1998;74:1556-65.
8. Trnini V, Baran i M, Nazor M. The five-factor model of personality and aggressiveness in prisoners and athletes. *Kinesiology* 2008;40:170-81.
9. Block JA. Contrarian view of the five-factor approach to personality description. *Psychol Bull* 1995;117:187-215.
10. Norman WT. Toward an adequate taxonomy of personality attributes. *J Abnorm Soc Psychol* 1963;66:574-83.
11. Pervin LA. *Persönlichkeitstheorien*. München: Ernst Reinhardt Verlag; 1993.
12. Costa PT Jr, McCrae RR. Stability and change in personality assessment: the revised neo personality inventory in the year 2000. *J Pers Assess* 1997;68:86-94.
13. Tsaousis I, Kerpelis P. The traits personality questionnaire 5 (Tpque5): the psychometric evaluation of a shortened version. *Eur J Psychol Assess* 2004;20:180-91.
14. Saucier G, Georgiades S, Tsaousis I, Goldberg RL. The factor structure of Greek personality adjectives. *J Pers Soc Psychol* 2005;88:856.
15. Goldberg LR, Johnson JA, Eber HW, et al. The international personality item pool and the future of public-domain personality measures. *J Res Pers* 2006;40:84-96.
16. Mlacic B, Goldberg LR. An analysis of a cross-cultural personality inventory: the IPIP Big-Five factor markers in Croatia. *J Pers Assess* 2007;88:168-77.
17. Socha A, Cooper CA, McCord DM. Confirmatory factor analysis of the M5-50: an implementation of the IPIP item set. *Psychol Assess* 2010;22:43-9.
18. Digman JM. Higher order factors of the big five. *J Pers Soc Psychol* 1997;73:1246-56.
19. Loevinger J. Measurement of personality: true nor false. *Psychol Inq* 1993;4:1-16.
20. Benet V, Waller NG. The big seven factor model of personality description: evidence for its cross-cultural generality in a Spanish sample. *J Pers Soc Psychol* 1995;69: 701-18.
21. Becker P, Kupsch M. The four-plus-X factor model for the description of normal and disordered personality as an alternative to the five factor model: a test of two models. *Trierer Psychologische Berichte* 2002;29.
22. Ashton MC, Lee K, Perugini M, et al. A six-factor structure of personality-descriptive adjectives: solutions from psycholexical studies in seven languages. *J Pers Soc Psychol* 2004;86:356-66.
23. Pervin L, Cervone D, Oliver J. *Personality: theory and research*. 9<sup>th</sup> ed. Hoboken: John Wiley & Sons; 1999.
24. Aghababaei N, Tabik MT. Gratitude and mental health: differences between religious and general gratitude in a Muslim context. *Ment Health Relig Cult* 2013;16:761-6.
25. Aghababaei N, Arji A. Well-being and the HEXACO model of personality. *Pers Individ Dif* 2014;56:139-42.
26. Cooper CA, Carpenter D, Reiner A, McCord DM. Personality and job satisfaction: evidence from a sample of street-level bureaucrats. *Int J Pub Adm* 2014;37:155-62.
27. Bradley BH, Klotz AC, Postlethwaite BE, Brown KG. Ready to rumble: how team personality composition and task conflict interact to improve performance. *J Appl Psychol* 2013;98:385-92.
28. Benotsch EG, Jeffers AJ, Snipes DJ, et al. The five factor model of personality and the non-medical use of prescription drugs: associations in a young adult sample. *Pers*

- Individ Dif 2013;55:725-868.
29. Chan CSR, Park HD. The influence of dispositional affect and cognition on venture investment portfolio concentration. *J Bus Venturing* 2013;28:397-412.
  30. Cooper CA, Knotts HG, McCord DM, Johnson A. Taking personality seriously: the five factor model and public management. *Am Rev Pub Adm* 2012;43:397-415.
  31. Gosling SD, Rentfrow PJ, Swann, WB. A very brief measure of the big-five personality domains. *J Res Pers* 2003;37:504-28.
  32. Lyrakos GN, Papazafiropoulou A, Bousmpoulas S, et al. The correlation of the personality in relation to the existence of diabetes mellitus. 1<sup>st</sup> National Congress of Mental Health in Primary Health Care, 10-13 October 2013, Athens.
  33. Muck PM, Hell B, Gosling SD. Construct validation of a short five-factor model instrument. A self-peer study on the German adaptation of the ten-item personality inventory (TIPI-G). *Eur J Psychol Assess* 2007;23:166-75.
  34. John OP, Srivastava S. The Big Five trait taxonomy: history, measurement, and theoretical perspectives. In: Pervin LA, John OP, eds. *Handbook of personality*. 2<sup>nd</sup> ed. New York: Guilford; 1999. pp 102-38.
  35. Medical Outcomes Trust. Trust introduces new translation criteria. *Trust Bull* 1997;5:1-4.
  36. Guillemin F, Bombardier C, Beaton D. Cross cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993;46:1417-32.
  37. Benet V, Waller NG. The Big Seven factor model of personality description; evidence for its cross-cultural generality in a Spanish sample. *J Pers Soc Psychol* 1995;69:701-18.
  38. Lyrakos G, Damigos D, Mavreas V, et al. A translation and validation study of the life orientation test-revised in the Greek-speaking population of nurses among three hospitals in Athens and Ioannina. *Soc Indic Res* 2010;95:129-42.
  39. West SG, Finch JF, Curran PJ. Structural equation models with nonnormal variables. In: Hoyle RH, ed. *Structural equation modeling: concepts, issues, and applications*. Thousand Oaks: Sage Publications; 1995. pp 56-75.
  40. Curran PJ, West SG, Finch JF. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychol Methods* 1996;1:16-29.
  41. Bollen KA. *Structural equations with latent variables*. New York: Wiley; 1989.
  42. Dragioti E, Vitoratou S, Kaltsouda A, et al. Psychometric properties and factor structures of the Greek version of the Cardiac Anxiety Questionnaire (CAQ). *Psych Rep* 2011;109:77-92.
  43. Jöreskog KG, Sörbom D. *LISREL-VI user's guide*. Mooresville: Scientific Software; 1984.
  44. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, eds. *Testing structural equation models*. Newbury Park: Sage; 1993. pp 136-162.
  45. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psych Bull* 1980; 88:588-606.
  46. Bentler PM. Comparative fit indexes in structural models. *Psych Bull* 1990;107: 238-46.
  47. Cronbach LJ. Coefficient alpha and the interval structure of tests. *Psychometrika* 1951;16:297-334.
  48. Kaiser HF. The application of electronic computers to factor analysis. *Edu Psychol Meas* 1960;20:141-51.
  49. Cattell RB. The scree test for the number of factors. *Multivariate Behav Res* 1966; 1:245-76.
  50. Clark LA, Watson D. Constructing validity: basic issues in objective scale development. *Psychol Assess* 1995;7:309-19.
  51. Nunnally JC. *Psychometric theory*. 2nd ed New York: McGraw Hill; 1978.
  52. Schönrock-Adema J, Heijne-Penninga M, van Hell EA, Cohen-Schotanus J. Necessary steps in factor analysis: enhancing validation studies of educational instruments. The PHEEM applied to clerks as an example. *Med Teach* 2009;31: e226-32.
  53. Zheng L, Goldberg LR, Zheng Y, et al. Reliability and concurrent validation of the IPIP Big-Five factor markers in China: consistencies in factor structure between internet-obtained heterosexual and homosexual samples. *Pers Individ Dif* 2008;45:649-54.
  54. Gow AJ, Whiteman MC, Pattie A, Deary IJ. Goldberg's "IPIP" Big-Five factor markers: internal consistency and concurrent validation in Scotland. *Pers Individ Dif* 2005;39:317-29.
  55. Borkenau P, Ostendorf F. Comparing exploratory and confirmatory factor analysis: a study of the five-factor model of personality. *Pers Individ Dif* 1990;11:515-24.
  56. McCrae RR, Zonderman AB, Costa PT Jr, et al. Evaluating replicability of factors in the revised NEO personality inventory: confirmatory factor analysis versus procrustes rotation. *J Pers Soc Psychol* 1996;70:552-66.
  57. Ehrhart KH, Roesch SC, Ehrhart MG, Kilian B. A test of the factor structure equivalence of the 50-item IPIP Five-factor model measure across gender and ethnic groups. *J Pers Assess* 2008;90:507-16.