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

Maria Ypofanti,1,2 Vasiliki Zisi,2 Nikolaos Zourbanos,2 Barbara Mouchtouri,3 Pothiti Tzanne,3 Yannis Theodorakis,2 Georgios Lyra,1,4

1Department of Psychiatry, General Hospital Nikaia-Piraeus Agios Panteleimon, Athens; 2Department of Physical Education and Sport Science University of Thessaly; 3Psychiatric Hospital (Daphni), Athens; 4Second 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=183) 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 (α=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.1,2 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),3 which was based on the notion that there are only five basic dimensions of personality.4 Although the NEO PI-R has been used as a tool in numerous studies,5,6 it was argued that the five factors personality model was an arbitrary selection procedure among interactions of experiential factors.7 The model was actually aiming to incorporated concepts related to an individual’s predisposition,7,10 12 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,13-15 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,15 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) extraversion-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. Macic and Goldberg weighed the questionnaire for the Croatian population,16 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.17

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,18 while others doubted that five factors are enough to perceive the total perspective of the personality function.19,20 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.21,22 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,23 the IPIP has been used in several studies. For example, it has been used in measuring religious gratitude and wellbeing,24-26 in examining the effects of personality traits on team processes and outcomes,27 on non-medical use of prescription drugs,28 even in exploring how an investor’s dispositional affect and cognitive style influence venture investment portfolio concentration.29 Finally, we should mention that public managers are aware of personality assessment, use it in their jobs, and are generally convinced of its efficacy.30

Since the IPIP’s inception, portions of the 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.31 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.

Correspondence: Georgios N. Lyra, General Hospital Nikaia, Piraeus “Agios Panteleimon”, Department of Psychiatry, D. Mandouvalou 3, Nikaia, 18454, Greece.
Tel: +30.2132.077146 - Fax: +30.2132.077143.
E-mail: geolyr@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).
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 postgraduate (5.4%) degrees. In terms of marital status, 49% were married, (44.8%) were single, (5.4%) were divorced and the remainder 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.15 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.1 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).

The internal consistency of the factors (Cronbach α) reported by Goldberg,15 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-3: 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),31 which was also translated into Greek following the same way forward and backward translation and was validated and used in a previous study with diabetic patients.32 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 α 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.33 The convergent validity for the TIPI with the 44-item BFI was comparable to the convergent validity of other longer multi-item FFM measures.34 Discriminant correlations were substantially lower than convergent correlations and the TIPI demonstrated good stability as indexed by test-retest correlations.35

Translation of the instrument

The translation strategy was based on minimal criteria developed by the Scientific Advisory Committee of the Medical Outcomes Trust.35 Translation was performed using the multiple forward and backward translation protocol recommended by Guillemin.36 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 working in Greece for the last 10 years) who were blinded to the original version retranslated the re-collated Greek version into the source language (back translation) which is the recommended procedure for creating semantic equivalence.37

The last step of the translation procedure was the pre-testing of the translated instrument in a small population of hospital stuff following a cognitive debriefing process as it is explained by Lyrakos et al.38 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 (≥20%), except for items where high rates in this response category were expected, ceiling, and floor effects (≥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 (skewness2, kurtosis7) and still within acceptable values for normality.39,40

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,41 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 Dragiotti et al.’s methodology for CFA42 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,43 the relative chi square (χ²/df)43 is also provided. Three further absolute meas-
ures of fit are reported, namely: goodness-of-fit index (GFI), Bollen’s relative fit index (RFI) and root mean square error of approximation (RMSEA). 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 (NFI) and the comparative fit index (CFI); ii) internal consistency reliability of the instrument was assessed using Cronbach’s α coefficient (1951), 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 (TIP). 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;19 ii) Percentage of variance accounted for by the retained component(s); and iii) Screeplot. 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.51,52

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±SD: 0.89±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 single, 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 TIP and the retet 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(E1411) (r=0.002), and E-36(X68) (r=0.100). Finally item E-46(H661) was found to have a negative loading even thought it was supposed to be positive and was decided to be omitted from the questionnaire (r =-0.146).

Then, Cronbach’s α 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 construction 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-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 α) 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 α, 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 α for this subset at the initial test period was 0.88, and was reduced by 0.05 at the retest period (a=0.83). Correlations between the test and retest mean scores of these participants were found to be above 0.3. Only the factor with 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).

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 Intelect factor had a higher correlation with the Agreeableness r=0.322 (P<0.001) and the Emotional stability items, with a smaller, but >0.40 loading in the same factor, while E-36(X68) had its highest loading with the intellect items.

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 (χ²=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 that 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 Schinnock-Adema criteria for factor analysis).44

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 that the five factor model.44 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 r>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 Intelect 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-36(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-36(X68) 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.\textsuperscript{15-17} The relative and absolute fit measures indicated that the five-factor model fit the data inadequately but was more acceptable that the independent and the model with no intercorrelations among the latent variables (Table 4). Chi square was statistically 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.\textsuperscript{43} The values of RFI (0.63) and RMSEA (0.058) also

Table 1. Means and standard deviation of the questionnaires.

<table>
<thead>
<tr>
<th>N.</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
<th>Skewness Mean</th>
<th>SE</th>
<th>Mean</th>
<th>Kurtosis SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>520</td>
<td>18</td>
<td>77</td>
<td>37.41</td>
<td>11.29</td>
<td>0.571</td>
<td>0.107</td>
<td>0.020</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>519</td>
<td>11</td>
<td>64</td>
<td>38.71</td>
<td>8.98</td>
<td>-0.684</td>
<td>0.107</td>
<td>-0.174</td>
</tr>
<tr>
<td>Emotional stability</td>
<td>519</td>
<td>10</td>
<td>50</td>
<td>29.45</td>
<td>8.44</td>
<td>-0.022</td>
<td>0.107</td>
<td>-0.534</td>
</tr>
<tr>
<td>Intellect</td>
<td>519</td>
<td>17</td>
<td>50</td>
<td>38.62</td>
<td>6.15</td>
<td>-0.445</td>
<td>0.107</td>
<td>0.341</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>518</td>
<td>21</td>
<td>52</td>
<td>42.82</td>
<td>5.50</td>
<td>-0.936</td>
<td>0.107</td>
<td>0.366</td>
</tr>
<tr>
<td>Conscientiousness retest</td>
<td>194</td>
<td>1</td>
<td>7</td>
<td>5.72</td>
<td>1.25</td>
<td>-0.233</td>
<td>0.291</td>
<td>0.088</td>
</tr>
<tr>
<td>Emotional stability retest</td>
<td>68</td>
<td>10</td>
<td>45</td>
<td>29.68</td>
<td>7.27</td>
<td>-0.398</td>
<td>0.291</td>
<td>0.032</td>
</tr>
<tr>
<td>Intellect retest</td>
<td>68</td>
<td>29</td>
<td>50</td>
<td>39.96</td>
<td>4.58</td>
<td>-0.032</td>
<td>0.291</td>
<td>-0.332</td>
</tr>
<tr>
<td>Agreeableness retest</td>
<td>68</td>
<td>31</td>
<td>50</td>
<td>43.54</td>
<td>6.63</td>
<td>-0.172</td>
<td>0.291</td>
<td>0.086</td>
</tr>
<tr>
<td>Extraversion retest</td>
<td>68</td>
<td>17</td>
<td>50</td>
<td>37.38</td>
<td>7.01</td>
<td>-0.233</td>
<td>0.291</td>
<td>0.175</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Cronbach $\alpha$</th>
<th>60 items</th>
<th>Cronbach $\alpha$</th>
<th>50 items</th>
<th>Range of item total correlations, $r$</th>
<th>Omitted items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conscientiousness</td>
<td>0.856</td>
<td>0.875</td>
<td>0.403-0.725</td>
<td>C3-H250, C13-H1362</td>
<td></td>
</tr>
<tr>
<td>Emotional stability</td>
<td>0.836</td>
<td>0.849</td>
<td>0.416-0.692</td>
<td>ES9-E141, ES19-X156</td>
<td></td>
</tr>
<tr>
<td>Intellect</td>
<td>0.762</td>
<td>0.780</td>
<td>0.353-0.519</td>
<td>E10-X176, I20-H1230</td>
<td></td>
</tr>
<tr>
<td>Agreeableness</td>
<td>0.702</td>
<td>0.758</td>
<td>0.210-0.573</td>
<td>A7-H21, A7-H107</td>
<td></td>
</tr>
<tr>
<td>Extraversion</td>
<td>0.754</td>
<td>0.791</td>
<td>0.359-0.555</td>
<td>E36-X68, E46-H661</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.866</td>
<td>0.882</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

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

*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). The values of the relative fit measures were moderate (NNFI=0.67, CFI=0.73); values higher than 0.90 indicate close fit. Since the above results suggested the initial model did not fit the data adequately in χ², 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 α 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 correlation 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 (Intelect 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 revel 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 Bonferroni 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 (\( M=36.7\pm9.4 \)) and married people (\( M=40.5\pm8.1 \)) 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 (\( M=30.9\pm8.5 \)) and people with postgraduate degrees (\( M=35.6\pm7.5 \)). Also, significant differences were found in Emotional Stability between people with 9 years of compulsory education (\( M=24.2\pm9.2 \)) and people with 12 years of education (\( M=29.2\pm8.4 \)) (MD= -5.01, \( P=0.49 \)), as well as with undergraduates (\( M=29.8\pm8.2 \)) (MD= -5.63, \( P=0.01 \)) and Master graduates (\( M=32.7\pm7.9 \)) (MD= -8.50, \( P=0.004 \)) respectively. Finally according to residency, independent sample t test shown a significant difference between residents of Athens (\( M=34.9\pm7.8 \)) and providence (\( M=37.8\pm6.2 \)) (t= -3.766, \( P<0.001 \)) in Extraversion.

Regarding the age of the sample, Pearson r correlations were also calculates with IPIP and TIPI. There was a significant but low correlation in the Consciousness 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.1 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.

<table>
<thead>
<tr>
<th>Models</th>
<th>X²</th>
<th>Df</th>
<th>X²/df</th>
<th>RFI</th>
<th>RMSEA (P-close)*</th>
<th>NNFI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 items with intercorrelations</td>
<td>3223.20</td>
<td>1165</td>
<td>2.767</td>
<td>0.636</td>
<td>0.058 (0.000)</td>
<td>0.667</td>
<td>0.755</td>
</tr>
<tr>
<td>Independent model</td>
<td>9684.40</td>
<td>1275</td>
<td>7.596</td>
<td>0.000</td>
<td>0.113 (0.000)</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>50 items with no intercorrelations</td>
<td>3560.88</td>
<td>1175</td>
<td>3.031</td>
<td>0.601</td>
<td>0.063 (0.000)</td>
<td>0.632</td>
<td>0.716</td>
</tr>
<tr>
<td>47 items with all intercorrelations</td>
<td>3531.63</td>
<td>1024</td>
<td>3.449</td>
<td>0.687</td>
<td>0.055 (0.000)</td>
<td>0.715</td>
<td>0.778</td>
</tr>
<tr>
<td>47 items (with p&lt;2 in F3)</td>
<td>3551.83</td>
<td>1024</td>
<td>3.469</td>
<td>0.685</td>
<td>0.055 (0.000)</td>
<td>0.714</td>
<td>0.776</td>
</tr>
<tr>
<td>47 items with no intercorrelations</td>
<td>3937.63</td>
<td>1034</td>
<td>3.808</td>
<td>0.654</td>
<td>0.059 (0.000)</td>
<td>0.683</td>
<td>0.734</td>
</tr>
</tbody>
</table>

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

Table 5. Independent sample t-test for sex.

<table>
<thead>
<tr>
<th></th>
<th>Males (N=281)</th>
<th>Females (N=530)</th>
<th>t</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>37.17</td>
<td>9.39</td>
<td>39.56</td>
<td>8.63</td>
</tr>
<tr>
<td>Emotional stability</td>
<td>31.02</td>
<td>8.51</td>
<td>28.58</td>
<td>8.29</td>
</tr>
<tr>
<td>Intellect</td>
<td>38.99</td>
<td>6.07</td>
<td>38.42</td>
<td>6.20</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>40.82</td>
<td>5.76</td>
<td>43.94</td>
<td>5.03</td>
</tr>
<tr>
<td>Extraversion</td>
<td>35.91</td>
<td>7.31</td>
<td>34.99</td>
<td>7.75</td>
</tr>
</tbody>
</table>

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, while in the Scottish version, 6 of the Agreeableness items loaded together, with a further 3 loading highest with the Extraversion items.

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.6,57

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, 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.16

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 criterion 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, as well as Ehrrhart et al., 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 rise 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 thought 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