The 16PF5 and the NEO-PI-R in Spanish and Swiss Samples: A Cross-Cultural Comparison

Anton Aluja1, Jérôme Rossier2, Luis F. Garcia1, and Sabrina Verardi2

1University of Lleida, Spain, 2University of Lausanne, Switzerland

Abstract. This study compared the Spanish (Castilian) and French versions of the 16PF5 and of the NEO-PI-R in Spanish and Swiss samples. The five-factor solution for the 16PF5 only seems clear for the Castilian version, but not for the French version. Indeed, the congruence coefficients for the Tough-Mindedness and the Self-Control dimensions are low. On the other hand, the five-factor solutions are highly similar for both countries concerning the NEO-PI-R, and the congruence coefficients are above .95 for all five dimensions. The low cross-cultural replicability for the 16PF5 makes it difficult to analyze the differences at the mean level for this inventory. For the NEO-PI-R, the differences are generally very small and globally account for 2.6% of the total variance. Spaniards seem to have slightly lower scores on Actions and slightly higher scores on Dutifulness. These differences could either be due to translation problems, sample selection, or cultural differences.

Keywords: 16PF5, NEO-PI-R, five-factor model, effect size, congruence coefficients

The structural validity of the Five-Factor Model (FFM) of personality has been solidly established and seems now the most common dimensional approach to personality traits (McCrae & Allik, 2002). The fifth edition of the 16PF (16PF5) has demonstrated a stable second order structure in the English, French, and Spanish languages (Aluja & Blanch, 2003; Conn & Rieke, 1994; Mogenet & Rolland, 1995). The revised NEO Personality Inventory (NEO-PI-R) has also demonstrated a robust and stable structure in different cultures (McCrae & Allik, 2002; Rossier, Dahourou, & McCrae, 2005), and also in the Spanish and French languages (Aluja, Garcia, Garcia, & Seisdedos, in press; Rossier, Wenger, & Berthoud, 2001). Cattell’s model and the FFM of personality have a common origin: The study of the taxonomy of personality based on the lexical approach, and both models propose a hierarchical structure of personality.

The 16PF questionnaire has been translated into more than 40 languages, and it is probably one of the most widely used personality questionnaires. In spite of some historical critics (Eysenck, 1986), the 16-factor structure has been considered as relatively robust (Hofer, Horn, & Eber, 1997), and the connection between the primary and secondary dimension was improved in the fifth edition (Aluja & Blanch, 2004; Chernyshenko, Stark, & Chan, 2001). Conn and Rieke (1994) proposed a second order structure based on five dimensions: Extraversion (Ex), Anxiety (An), Tough-Mindedness (Tm), Independence (In), and Self-Control (Sc). These dimensions do not coincide exactly with the dimensions of the NEO-PI-R: Extraversion (E), Neuroticism (N), Openness (O), Agreeableness (A), and Conscientiousness (C). Nevertheless, there is a good correspondence between An and N, Sc and C, Tm and O, and between the two scales of Extraversion (respectively Ex for the 16PF5 and E for the NEO-PI-R). A poor correspondence was observed between In and A (Byravan & Ramanaiah, 1995; Conn & Rieke, 1994; Rossier, Meyer de Stadelhofen, & Berthoud, 2004).

The cross-cultural approach offers a unique way of establishing the validity of personality models. Following the emergence of the FFM (Digman, 1990; Peabody & Goldberg, 1989) a lot of attention has been paid to cross-cultural studies, particularly with the NEO-PI-R (McCrae & Allik, 2002; Rossier, Dahourou, & McCrae, 2005). The goal in these studies was to demonstrate that the features of personality vary only a little as a function of culture. In a recent review of Rolland (2002) it is asserted that the cross-cultural generalizability of the FFM is clearly evident, although certain facets of Extraversion and Agreeableness seem to be more sensitive to cultural context. The methodology used to compare factorial structures obtained with the NEO-PI-R in different cultures is the comparison of factor matrices and the calcul-
lation of congruence coefficients (McCrae, Zonderman, Costa, Bond, & Paunonen, 1996). This method can be used to compare the structures underlying personality, but high congruence coefficients can be obtained from samples with substantial differences at the means score level. For example, in Spain the standardization of the 16PF5 and of the NEO-PI-R were made with nonanonymous samples that completed the questionnaires in a selection situation (see Spanish manuals by Russell & Karol, 1995; Costa & McCrae, 1999). The means obtained for these samples differ substantially from the means obtained with these two instruments in anonymous and voluntary groups. The tendency is for subjects in a selection situation to obtain higher scores on dimensions related to social adaptation and lower scores on dimensions related to emotional instability, dominance, and impulsiveness (Aluja & Blanch, 2002). When comparing the means in independent groups, it is necessary to keep in mind the effect sizes and the power analysis (Cohen, 1962; 1988). Statistical power refers to the probability of not making a Type II error. The effect size tells us something very different from the p-value, which indicates the obtained probability of Type II error in a test of statistical significance. According to Rosenthal (1994), a p-value reported as "statistically significant" does not mean that the effect is large, nor does a p-value reported as "not significant" imply a trivial result. In the present study we compared the Spanish and Swiss scores obtained on the primary factors of the 16PF5 and on the facets of the NEO-PI-R and calculated the effect sizes.

It was expected that the structures underlying both inventories would be similar in both cultures. Concerning the mean differences, the pattern of differences should be similar with both personality inventories. Indeed, differences should not depend on the measurement method. Nevertheless, this hypothesis can be tested only if both inventories have a high cross-cultural replicability, as expected. Thus, the aim of the present work was to study the cross-cultural similarities and differences between the French and Spanish versions of the 16PF5 and the NEO-PI-R.

Method

Subjects
The Spanish participants were 636 undergraduate students, as well as their friends and relatives, who anonymously completed the 16PF5 together with other personality questionnaires (348 women and 287 men; mean age = 25.1, SD = 9.2). In another independent study, participants were 1,006 Spanish undergraduate students, who anonymously completed the NEO-PI-R (639 women and 367 men; mean age = 22.3, SD = 5.0). The Swiss samples for both inventories were broadly representative of the French-speaking Swiss population and made up of 386 participants for the 16PF5 (230 woman and 156 men; mean age = 32.5, SD = 13.4) and 1090 for the NEO-PI-R (620 women and 470 men; mean age = 34.1, SD = 14.6). All subjects participated voluntarily and anonymously in this study.

Measures

16PF
The 16PF5 is a self-rating questionnaire of 185 questions. For the Spanish sample, the fifth edition of the 16PF, adapted from the original English version of Cattell, Cattell, and Cattell (1993) was used. The Spanish version has the same number of items as the original version, though in a different order (Russell & Karol, 1995). Internal consistencies for the 16 factors range from .54 to .84, which are equivalent to those reported by Conn and Rieke (1994) and Aluja and Blanch (2002). The French version of the 16PF5, which was adapted from the original English version by Mogenet and Rolland (1995), was used with the Swiss sample. Internal consistencies range from .56 to .84.

NEO-PI-R
The NEO-PI-R is a well-known measure of the FFM of personality (Costa & McCrae, 1992). The Spanish version of the NEO-PI-R used in this study has acceptable psychometric properties, which are similar to the original version (Aluja, García, & García, 2002, 2003; Aluja, García, García, & Seisdedos, in press). The French NEO-PI-R version also obtained good psychometric properties in France (Rolland & Petot, 1998) and in the French-speaking part of Switzerland (Rossier, Wenger, & Berthoud, 2001; Rossier, Meyer de Stadelhofen, & Berthoud, 2004).

Analysis
Principal Components (PC) does not distinguish between common, unique, and error variance, and the components extracted cannot be equated to common factors. If the goal is to arrive at a parsimonious representation of the associations among measured variables, the PC is not appropriate (Fabrigar, Wegener, Maccallum, & Strahan, 1999). In these cases, EFA extraction methods, such as Principal Axis (PA), are strongly recommended. This recommendation is reinforced when communalities are low (< 0.40), a pattern usually found in factor analysis at
the item level. Also, most of the studies have performed an orthogonal rotation. Since psychological variables usually correlate, oblique rotation methods are preferable to orthogonal ones (Fabrigar et al., 1999). Following these rules, a PA extraction with direct oblimin rotation procedure was conducted in all cases.

After a factorial analysis for each inventory and sample, we computed the coefficients of congruence for each inventory in order to check the invariance according to culture. The coefficients of congruence were also calculated for each variable. The congruence coefficient \( r_c \) is an index of factor similarity (Cattell, 1978). The congruence is considered high if \( r_c > .90 \), borderline if \( .90 \geq r_c > .80 \), and poor if \( r_c \leq .80 \). This method allows comparing factorial structures in different samples (McCrae, Zonderman, Costa, Bond, & Paunonen, 1996).

In order to assess mean score differences by country and gender we conducted two-way ANCOVAs with age as a covariate. Indeed, the Spaniard and Swiss samples differed substantially in age, which is known to be related to personality trait levels. This relation seems stable across culture as stated by McCrae, Costa, Lima, and colleagues (1999). Nevertheless, as both samples were rather large, it is possible that the F-statistic offers significant differences starting from very small differences among the means. For this reason, we also calculated the effect size \( \eta^2 \). \( \eta^2 \) gives an indication of the proportion of variance. \( \eta^2 \) will also be transformed in \( d \)-values (Cohen, 1988, pp. 273–288). \( d = .20 \) defines a small effect size, \( d = .50 \) defines a medium effect size, and \( d = .80 \) defines a large effect size.

### Results

#### 16PF Fifth Edition

For the two samples a principal axis analysis with oblique rotation on the 15 first order factors (excluding Factor B), forced to 5 factors, was conducted (Table 1). For the Spanish sample, the Kaiser-Meyer-Olkin measure of the sample adequacy was .72, with a Bartlett test of sphericity yielding a \( \chi^2 \) of 2420.91 (105 df; \( p < .001 \)). The five factors obtained eigenvalues higher than 1 and explain 65.04% of the variance before rotation (Factor I: 20.64; Factor II: 16.64%; Factor III: 11.17%; Factor IV:,

---

### Table 1: 16PF5 – Principal axis analysis with direct oblimin rotation, congruence coefficients, and correlations between factors for the Spanish and the Swiss samples.

| | Spanish sample | | | | Swiss sample | | | |
| | Ex | An | Sc | In | Tm | Ex | An | Sc | In | Tm | CC |
| A | .67 | .07 | .17 | −.03 | .25 | .39 | .21 | .14 | .16 | −.29 | .99 |
| C | .02 | −.56 | .11 | .33 | −.11 | .06 | −.76 | .18 | .04 | −.13 | .93 |
| E | .02 | .18 | .10 | .81 | .05 | −.03 | −.06 | −.01 | .77 | .07 | .99 |
| F | .60 | .02 | −.27 | .15 | .07 | .84 | −.08 | −.23 | −.01 | .25 | .91 |
| G | −.09 | −.15 | .62 | −.04 | −.05 | .00 | .04 | .63 | −.10 | −.03 | .81 |
| H | .32 | −.29 | −.06 | .40 | .11 | .35 | −.25 | −.04 | .46 | −.11 | .97 |
| I | .27 | .11 | .09 | −.20 | .46 | .12 | .43 | −.03 | −.21 | −.22 | .89 |
| L | −.14 | .53 | .00 | .09 | .01 | .03 | .25 | .08 | .12 | .70 | .86 |
| M | −.33 | .12 | −.29 | −.11 | .67 | .01 | .22 | −.65 | .02 | .13 | .82 |
| N | −.48 | .16 | .05 | −.05 | .04 | −.25 | −.19 | −.04 | −.33 | .43 | .57 |
| O | .18 | .48 | .23 | −.29 | .26 | −.02 | .61 | .08 | −.14 | .03 | .98 |
| Q1 | −.04 | −.15 | −.07 | .22 | .61 | .08 | −.02 | −.37 | .29 | −.24 | .94 |
| Q2 | −.74 | .03 | .00 | .06 | .13 | −.48 | .03 | −.06 | .04 | .09 | .98 |
| Q3 | .00 | .05 | .59 | .13 | −.03 | −.05 | .05 | .56 | .18 | .14 | .86 |
| Q4 | .00 | .70 | −.08 | .16 | −.11 | −.09 | .49 | −.01 | .22 | .10 | .74 |

Note. Ex = Extraversion; An = Anxiety; Sc = Self-Control; In = Independence; Tm = Tough-Mindedness; A = Warmth; C = Emotional Stability; E = Dominance; F = Liveliness; G = Rule-Consciousness; H = Social Boldness; I = Sensitivity; L = Vigilance; M = Abstractedness; N = Privateness; O = Apprehension; Q1 = Openness to Change; Q2 = Self-Reliance; Q3 = Perfectionism; Q4 = Tension.
9.08% and Factor V: 7.51%). For the Swiss sample, the Kaiser-Meyer-Olkin measure of the sample adequacy was of .76, with a Bartlett test of sphericity yielding a $\chi^2$ of 1388.93 (105 df, $p < .001$). The five factors also obtained eigenvalues higher than 1 and explain 63.34% of the variance before rotation (Factor I: 20.22%; Factor II: 15.00%; Factor III: 11.52%; Factor IV: 10.15%, and Factor V: 6.45%). Correlations between factors are shown in Table 1.

In the Spanish sample, the number of second order factors is clearly 5, while for the Swiss sample a four-factor solution seems more adequate according to the Scree test (see Figure 1). Concerning the cross-cultural replicability of the structure, we obtained congruence coefficients higher than .90. The congruence was borderline for Sc and poor for Tm. For the Spanish sample, the cross-cultural replicability of the structure was of .76, with a Bartlett test of sphericity yielding a $\chi^2$ of 1388.93 (105 df, $p < .001$). The five factors also obtained eigenvalues higher than 1 and explain 63.34% of the variance before rotation (Factor I: 20.22%; Factor II: 15.00%; Factor III: 11.52%; Factor IV: 10.15%, and Factor V: 6.45%). Correlations between factors are shown in Table 1.

Differences by country and gender were assessed with two-way ANCOVAs with age as a covariate. Table 2 shows means, standard errors for women and men, and for both samples. Age-adjusted means show the same patterns as the unadjusted means. Generally the differences are small. Considering the effect size, a large difference between the two countries was observed for Q1 (d = 1.26), a medium difference was observed for G (d = .76), L (d = .56), and Q2 (d = .70), and a small difference was observed for C (d = .40), F (d = .28), M (d = .24), O (d = .22), and Q3 (d = .40). Differences on the remaining primary factors were associated with a $d$ below .20. Thus, Spaniards score globally lower on Openness to change but higher on Rule-Consciousness, Vigilance, and Self-Reliance. The mean difference between the two countries is associated to a small effect size (d = .44). Concerning the sex differences, a large difference was observed only for I (d = 1.16), a medium difference was observed for A (d = .54) and O (d = .58), and a small difference was observed for B (d = .28), C (d = .30), E (d = .32), H (d = .24), N (d = .20), Q2 (d = .24), and Q4 (d = .30). Women score globally higher on Sensitivity, Warmth, and Openness to Change. The mean difference between women and men was associated to a small effect size (d = .36). Concerning the interaction between country and sex, only a small effect size was observed for A (d = .28). The mean effect size for the interaction did not reach a $d$-value of .20. The effect size associated to age was medium for F (d = .61) and small for B (d = .26), C (d = .30), G (d = .42), M (d = .22), Q1 (d = .29), Q2 (d = .21), and Q3 (d = .25).
NEO P-I-R

For the two samples, a principal axis analysis with direct oblimin rotation on the 30 facets, forced to 5 factors, was conducted (Table 3). For the Spanish sample, the Kaiser-Meyer-Olkin measure of the sample adequacy was of .86, with a Bartlett test of sphericity yielding a $\chi^2$ of 12731.92 ($435 \text{ df}$, $p < .001$). The five factors also obtained eigenvalues higher than 1 and explain 56.28% of the variance before rotation (Factor I: 18.85%; Factor II: 13.81%; Factor III: 9.78%; Factor IV: 8.10%, and Factor V: 7.74%). For the Spanish sample, the Kaiser-Meyer-Olkin measure of the sample adequacy was of .86, with a Bartlett test of sphericity yielding a $\chi^2$ of 14486.92 ($435 \text{ df}$, $p < .001$). The five factors also obtained eigenvalues higher than 1 and explain 58.08% of the variance before rotation (Factor I: 17.76%; Factor II: 13.95%; Factor III: 12.54%; Factor IV: 8.15%, and Factor V: 5.67%). With the criteria of the eigenvalues higher than or equal to one, six factors were obtained for both samples. Figure 1 shows the Scree plot for the Spanish and the Swiss samples. Also, correlations between factors are presented in Table 3.

In both factor matrices it can be seen that facets are well grouped in their respective factors with loadings higher than .40, except for E3 which loads .37 and .34 in the Spanish and Swiss samples. This facet loads negatively in the A and N factors in both matrices. Concerning the cross-cultural replicability, the total congruence coefficient was .97. For all five factors the congruence coefficient was above .95. For facet-scales the congruence coefficients ranged from .85 to 1 ($Mdn = .98$). The congruence coefficient was borderline for O4 and E4 and high for all other facet-scales.

Differences by country and gender were assessed with two-way ANCOVAs with age as a covariate. Table
Table 3. NEO-PI-R – Principal axis analysis with direct oblimin rotation, congruence coefficients, and correlations between factors for the Spanish and the Swiss samples.

<table>
<thead>
<tr>
<th></th>
<th>Spanish sample</th>
<th>Swiss sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>E</td>
</tr>
<tr>
<td>N1</td>
<td>−.83</td>
<td>.11</td>
</tr>
<tr>
<td>N2</td>
<td>−.67</td>
<td>−.07</td>
</tr>
<tr>
<td>N3</td>
<td>−.80</td>
<td>−.13</td>
</tr>
<tr>
<td>N4</td>
<td>−.63</td>
<td>−.24</td>
</tr>
<tr>
<td>N5</td>
<td>−.42</td>
<td>.31</td>
</tr>
<tr>
<td>N6</td>
<td>−.76</td>
<td>.00</td>
</tr>
<tr>
<td>E1</td>
<td>.11</td>
<td>.69</td>
</tr>
<tr>
<td>E2</td>
<td>−.06</td>
<td>.66</td>
</tr>
<tr>
<td>E3</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>E4</td>
<td>.03</td>
<td>.52</td>
</tr>
<tr>
<td>E5</td>
<td>.03</td>
<td>.34</td>
</tr>
<tr>
<td>E6</td>
<td>.28</td>
<td>.63</td>
</tr>
<tr>
<td>O1</td>
<td>−.07</td>
<td>.01</td>
</tr>
<tr>
<td>O2</td>
<td>−.15</td>
<td>.03</td>
</tr>
<tr>
<td>O3</td>
<td>−.23</td>
<td>.36</td>
</tr>
<tr>
<td>O4</td>
<td>.01</td>
<td>.07</td>
</tr>
<tr>
<td>O5</td>
<td>.06</td>
<td>−.21</td>
</tr>
<tr>
<td>O6</td>
<td>.17</td>
<td>.00</td>
</tr>
<tr>
<td>A1</td>
<td>.27</td>
<td>.33</td>
</tr>
<tr>
<td>A2</td>
<td>−.03</td>
<td>−.02</td>
</tr>
<tr>
<td>A3</td>
<td>.01</td>
<td>.38</td>
</tr>
<tr>
<td>A4</td>
<td>.13</td>
<td>.02</td>
</tr>
<tr>
<td>A5</td>
<td>−.17</td>
<td>−.02</td>
</tr>
<tr>
<td>A6</td>
<td>−.10</td>
<td>.15</td>
</tr>
<tr>
<td>C1</td>
<td>.31</td>
<td>.04</td>
</tr>
<tr>
<td>C2</td>
<td>−.14</td>
<td>−.01</td>
</tr>
<tr>
<td>C3</td>
<td>.10</td>
<td>.01</td>
</tr>
<tr>
<td>C4</td>
<td>−.13</td>
<td>.24</td>
</tr>
<tr>
<td>C5</td>
<td>.13</td>
<td>−.01</td>
</tr>
<tr>
<td>C6</td>
<td>.05</td>
<td>−.31</td>
</tr>
</tbody>
</table>

Congruence coefficients

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>E</th>
<th>O</th>
<th>A</th>
<th>C</th>
<th>N</th>
<th>E</th>
<th>O</th>
<th>A</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>.13</td>
<td>.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>.20</td>
<td>.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>.10</td>
<td>−.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>.03</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness; N1 = Anxiety; N2 = Anger Hostility; N3 = Depression; N4 = Self-Consciousness; N5 = Impulsivity; N6 = Vulnerability; E1 = Warmth; E2 = Gregariousness; E3 = Assertiveness; E4 = Activity; E5 = Excitement Seeking; E6 = Positive Emotions; O1 = Fantasy; O2 = Aesthetics; O3 = Feeling; O4 = Actions; O5 = Ideas; O6 = Values; A1 = Trust; A2 = Straightforwardness; A3 = Altruism; A4 = Compliance; A5 = Modesty; A6 = Tender-Mindedness; C1 = Competence; C2 = Order; C3 = Dutifulness; C4 = Achievement Striving; C5 = Self-Discipline; C6 = Deliberation.

4 shows means, standard errors for women and men, and for both samples. Age-adjusted means show the same patterns as the unadjusted means. Generally the differences are small. Countries account globally for less than 3% of the total variance and sex for about 1% of the variance. Considering the effect size, a medium difference between the two countries was only observed for O4 (d = .50) and C3 (d = .54), a small difference was observed for N2 (d = .36), N3 (d = .36), N4 (d = .22), N5 (d = .44), E6 (d = .34), O1 (d = .46), O2 (d = .28), O3 (d = .32), A1 (d = .24), A6 (d = .44), C1 (d = .40), C2 (d = .38), C4 (d = .48), C5 (d = .26), and C6 (d = .36). Differences on the remaining facet-scales were associated with a d below .20. Thus, Spaniards
score mainly lower on Action but higher on Dutifulness than Swiss. The mean difference between the two countries is associated to a small effect size ($d = .32$). Concerning the sex differences, a medium difference was observed only for N4 ($d = .30$) and N5 ($d = .32$) and C5 ($d = .22$). Women score mainly higher on Anxiety. The mean difference between women and men was associated to a small effect size ($d = .20$). Concerning the interaction between country and sex, only a small effect size was observed for N4 ($d = .28$) and C4 ($d = .22$). The mean effect size for the interaction did not reach a $d$-value of .20. The effect size associated to age was medium for E5 ($d = .61$) and small for N1 ($d = .29$), N2 ($d = .22$), N3 ($d = .31$), N4 ($d = .30$), N5 ($d = .40$), N6 ($d = .31$), E2 ($d = .29$), O1 ($d = .43$), O3 ($d = .40$), A1 ($d = .27$), A2 ($d = .33$), A4 ($d = .26$), C1 ($d = .20$), C3 ($d = .39$) and C5 ($d = .33$).
Discussion

With regard to the factor structures and the congruence coefficients, we obtained a borderline total congruence for the 16PF5. The factor congruence coefficient was poor for Sc and Tm. This is apparently because in the Swiss sample a clear five-factor structure was not found, and a four-factor structure could represent the observed data in a better way (Rossier et al., 2004). Subsequently, it may be hypothesized that the different sample sizes between Spanish and Swiss subjects could affect the second order factors structure, given that in a 1000-subjects French validation sample, the Tm factor was clearly identified (Mogenet & Roland, 1995). Indeed, Rossier et al. (2004) used the Horns method as a criterion for the factor extraction, and this method is sensitive to the number of variables or the sample size.

Bearing this in mind, for a sample of 386 and 15 variables, 4 factors would be advisable, while for a sample of 636 and also 15 variables five factors would be the better choice according to the eigenvalues obtained with the PA analysis on the 16PF5 for the Swiss sample (Horn, 1965; O’Connor, 2000). It is noteworthy that most studies capable of extracting a clear five-factor solution concerned large samples. Moreover, the idea behind the parallel analyses is to extract the factors that explain more than the factors extracted from random data. When the number of subjects is increased, it should be possible to distinguish factors that explain a smaller amount of the variance. Concerning the primary factors of the 16PF5, Privateness and Tension obtained very poor congruence coefficients. Five primary factors obtained a borderline value: Rule-Consciousness, Sensitivity, Vigilance, Abstractedness, and Self-Reliance. Only eight primary factors out of the 15 considered obtained a high congruence coefficient. Also, the congruence coefficients can be considered as an omnibus test of factor invariance. Analyzing specific models for each factor would be helpful when some of the factors are initially comparable and some were not. Confirmatory factor analysis permits these specific models to be developed and is recommended as the following step in studies about cross-cultural structural invariance (Aluja, et al., 2003; Aluja, Rossier, García, Angleitner, Kuhlman, & Zuckerman, submitted).

In the case of the NEO-PI-R, the congruence coefficients for the domains were always higher than .95. The total congruence was .97. Twenty-eight out of the 30 facets were associated with a high congruence coefficient. Only two facet-scales, Activity and Actions, obtained a borderline congruence coefficient. To summarize, the cross-language and cross-cultural replicability is high in the case of the NEO-PI-R, but poor in the case of the 16PF5.

The comparison of scores obtained by subjects of two different cultures raises the question of what is really measured. This question is particularly sensitive for data obtained from self-reports. When subjects are assessing and describing themselves, how do they really go about it? Even if personality has been reconceptualized as endogenous basic tendencies (McCrae, 2001), self-assessment procedures could be influenced by national stereotypes. In our case, it is difficult to tell if the differences obtained between the Spanish and the Swiss samples are due to translation problems, endogenous differences, or cultural differences (Heine, Lehman, Peng, & Greenholtz, 2002). It should be noted that structure equivalence is a condition for comparing groups, which does not exclude the possibility of shifts of the scores in one cultural group (Van de Vijver & Leung, 2001). Nevertheless, in our study the differences between both countries only affect some of the scales. If this is due to such a shift, the question remains as to why such a phenomenon only affects specific dimensions of personality. Also, note that the Spanish and Swiss versions do not contain the same set of translated items, given that both versions were adapted directly from the original English version. Therefore, some of the reported differences could be due to the slight differences between both versions. A strategy to contrast the different role of the facets in the means differences across countries is to test a model of “strong factorial invariance” (Hofer et al., 1997; Meredith, 1993). Since the exploratory factor analysis performed in most studies of personality structure does not consider this level of invariance, such an analysis would be desirable in future studies.

The analysis of means for the 16PF5 factors and for the NEO-PI-R facets in the Spanish and Swiss samples indicates that the global difference between the two countries is small. Note that the country variable accounts for less than 5% for the 16PF5 and for less than 3% for the NEO-PI-R, in spite of the sociodemographic differences between both samples. Also, results do not show important changes when the age variable is controlled. A large difference was only observed, according to the effect size, for one primary scale of the 16PF5, that is, Q1. In the present study, according to the 16PF5, Swiss respondents would be more liberal and more interested in change (Q1+), more self-sufficient (Q2+), less confident in others (L–), and informal (G–) than Spanish respondents, while the latter would be more extroverted than the former (Q2–). The result obtained on Q1 is supported by the fact that French and French-speaking Swiss have very high scores on Openness compared to most countries (Rossier et al., 2005). Moreover, for Q1 Swiss subjects obtain a very similar mean to that reported in the French manual (20.73 vs. 20.62). The mean score of the French and Swiss population is 3 points below the one

reported for the American population and 6 points below
the one reported for the Spanish population. For Q2
(9.53), the Swiss subjects score a little bit higher than the
Americans (8.35), although they have lower scores than
the French (7.57) and the Spanish (5.79) subjects (Mo-
genet & Rolland, 1995; Conn & Rieke, 1994; Russell &
Karol, 1995).

For the NEO-PI-R, the differences between the two
samples for the facets are generally less important as
measured by the effect size index. We observed no large
difference between the Swiss and the Spanish samples.
A medium effect size was only observed for Actions (O4)
and Dutifulness (C3), which corroborates the result ob-
tained with the 16PF5 for Rule-Consciousness (G). The
Spanish tended to score lower on the Openness facets,
especially on Fantasy (O2) and Actions (O4). This out-
come corroborates the result obtained with the 16PF5,
bearing in mind that Openness to change (Q1) is related
to Openness (O; Cattell & Cattell, 1995). Overall, the
effect sizes found for the differences on the dimensions
of both inventories are generally quite small. Moreover,
the Spanish and Swiss samples cannot be considered as
being totally representative of the general population.
Concerning the gender differences and the differences
according to age, we observed a similar pattern to the one
observed by Costa, Terracciano, and McCrae (2001). For
example, women obtained higher scores on Anxiety
(N1), Vulnerability (N6), Straightforwardness (A2), and
Altruism (A3). Concerning the age differences, we also
observed the same pattern of results as the one reported
by McCrae et al. (1999). For example, age had an influ-
ence on Excitement Seeking (E5).

The differences observed between the two countries
with the 16PF5 are slightly larger compared to those ob-
tained with the NEO-PI-R. This could be due to the poor
cross-cultural replicability of the 16PF5. For this reason,
we analyzed if there is any correlation between congru-
ence coefficients and effect sizes, but found no associ-
ation. The poor replicability of the 16PF5 makes it diffi-
cult to take the differences observed between the two
countries into consideration, whereas the very high rep-
licability of the NEO-PI-R makes it possible to be more
confident about the differences observed between the
countries. Concerning the NEO-PI-R, the differenc-
es were usually very small, but were medium for two
facet-scales. In this case, it would be unlikely that a bias
specifically affects the scores on these two scales.

Summing up, it can be said that the results with the
16PF5 indicate differences in the second order factor
structures that could be due to both sample size and slight
lexical differences in the formulation of items. The factor
structures of the NEO-PI-R obtained from the Spanish
and Swiss samples are highly similar, and the congru-
ence coefficients show a high cultural invariance, as ex-
pected. In the analyses of the differences between both
samples, according to the scores for each dimension, the
F-statistic provides biased estimates, thus, it is essential
to assess the effect size. It should also be highlighted that
mean differences are proportionally low, which could be
due to cultural differences between the two samples or to
translation differences in the respective languages.

References

del 16PF-5 en muestras americanas y españolas [Descriptive
and comparative analysis of the 16PF5 in American and Span-
ish samples]. Boletín de Psicología, 74, 27–38.

segundo orden del 16PF-5 en muestras americanas y españolas
[Replicability of 16PF5 second-order factors in American and

factors: An analysis of three parcelling methods. Personality
and Individual Differences, 37, 667–677.

of Zuckerman’s three structural models for personality through
the NEO-PI-R, ZKPQ-III-R, EPO-RS and Goldberg’s 50-bipolar
adjectives. Personality and Individual Differences, 33,
713–725.

Aluja, A., García, O., & García, L.F. (2003). Psychometric prop-
erties of the Zuckerman-Kuhlman Personality Inventory
(ZKPQ-III-R): A study of a shortened form. Personality
and Individual Differences, 34, 1083–1097.

Invariance of the “NEO-PI-R” factor structure across explor-
atory and confirmatory factor analyses. Personality and Indi-
vidual Differences.

Aluja, A., Rossier, J., García, L.F., Angleitner, A., Kuhlman, M.,
of the ZKPQ (ZKPQ-50-CC) adapted to English, French, Ger-
aman, and Spanish languages.

fifth edition from the perspective of the five-factor model. Psy-
chological Reports, 76, 555–560.

Cattell, R.B. (1978). The scientific use of factor analysis in behav-

the new fifth edition of the 16PF. Educational and Psycholog-
ical Measurement, 55, 926–937.

sonality Factor Questionnaire, fifth edition. Champaign, IL:
Institute of Personality and Ability Testing, Inc.

the hierarchical factor structure of the fifth edition of the 16PF:
An application of the Schmid-Leiman orthogonalization pro-

Cohen, L. (1962). The statistical power of abnormal-social psy-
chological research. Journal of Abnormal and Social Psychol-

© 2005 Hogrefe & Huber Publishers


Address for correspondence:

A. Aluja
Department of Pedagogy and Psychology
Faculty of Education Science
University of Lleida
Complex de la Caparrella, s/n.
E-25192 Lleida (Catalonia)
Spain
Tel. +34 973 702326
Fax +34 973 702305
E-mail aluja@pip.udl.es