Abstract. The present study examined the psychometric properties of the TEA Personality Test (TPT) in a sample of 23,062 Spanish adults. The TPT is a self-report questionnaire to be answered using a four-point Likert scale to assess personality at work and it is the third most frequently used tool by work psychologists in Spain. The reliability and validity analyses indicated that the TPT has adequate psychometric properties for the Spanish sample analyzed. Ordinal \( \alpha \) was used to calculate the internal consistency reliability of the scales. Results were higher than those of Cronbach's \( \alpha \) reported in the TPT's technical manual (\( p < .001 \)). Results from confirmatory factor analyses showed an acceptable goodness-of-fit for the theoretical three factors of the TPT's work personality model reported in the technical manual. The findings support the reliability and construct validity of the TPT.

Keywords: TPT, TEA personality test, reliability, construct validity

Introduction

In recent years, there has been a growing tendency in the literature to examine the psychometric properties of psychological tests, such being the topic of many research studies (v.g., Lenzenweger, Clarkin, Kernberg, & Foelsch, 2001; Tsaousis & Georgiades, 2009; Tsaousis & Kepelis, 2004). Spanish adaptations of tests have also emerged (v.g., Alonso-Arbiol, Balluerka, Shaver, & Gillath, 2008; Buela-Casal, Carretero-Dios, De los Santos-Roig, & Bermúdez, 2003; García, Aluja, & García, 2004; Gomà-i-Freixanet, Valero, Punti, & Zuckerman, 2004; Méndez, Hidalgo, & Ingélis, 2002; Penley, Wiebe, & Nwosu, 2003; Rintala et al., 2002; Rodríguez Fornells, Lorenzo-Seva, & Andrés-Pueyo, 2001; Ruipérez, Ibáñez, Lorente, Moro, & Ortet, 2001). These studies provide relevant scientific evidence to update psychometric data and validity of tests over the years.

In most cases, the tests’ technical manuals provide the most relevant information for their use and interpretation. However, authors and test developers emphasize the need to continue updating the psychometric properties and extending the research on which the test was based during its development and standardization (Arribas, Corral, & Pereña, 2010; Fernandez-Seara, Seisdedos, & Mielgo, 2008; Seisdedos, 2002a, 2002b, 2004a, 2004b, 2008). A systematic review of psychological tests and their properties is more widespread in the US than in Europe, and includes regular publications with the sole purpose of reviewing published tests (Geisinger, Spies, Carlson, & Plake, 2007).

Recently Muñiz and Fernandez-Hermida (2010) conducted a study on the most frequently used tests in Spain, using a questionnaire developed by the European Federation of Psychologists’ Associations (EFPA). Personality questionnaires most frequently used by work psychologists were, in this order: 16PF-5 (Cattell, Cattell, & Cattell, 2005), PAPI (Kostick, 1976), TPT (Corral, Pereña, Pamos, & Seisdedos, 2002), IPV (ECPA, 2005), MMPI-2 (Hathaway & McKinley, 2002), BFQ (Caprara, Barbaranelli, & Borgogni, 2005), MCMII-III (Millon, 2007) and NEO PI-R (Costa & McCrae, 2008). It also highlights SOSIA (Gordon, ECPA, & TEA Ediciones, 2008), BIP (Hossiep & Paschen, 2008) and compeTEA (Arribas & Pereña, 2009).

The TEA Personality Test (TPT), as the third most frequently used personality questionnaire by work psychologists in Spain, has no more psychometric studies than those reported in the original edition and the third revised edition. The technical manual reports the test’s reliability in terms of Cronbach’s \( \alpha \), with values ranging from 0.50 to 0.78 (median = 0.62). Construct validity was studied using an exploratory factor analysis (EFA), and its results suggest 3 second-order factors: Emotional Stability, Mental Openness, and Responsibility. There is no more information about reliability, construct validity, factorial solution fitness, or confirmatory theoretical structure in the manual.

Since the development of the TPT, the study of psychometric properties according to classical test theory has been significantly improved. In the field of reliability, Cron-
bach’s (1951) \( \alpha \) coefficient is the norm-referenced technique that many researchers have been using to measure the internal consistency of tests (Cortina, 1993; Zumbo & Rupp, 2004). However, Cronbach’s \( \alpha \) is based on the assumption of continuity of the variables and this assumption is not met by ordinal response items or Likert scales (Elosúa & Zumbo, 2008). Several studies provide evidence that Cronbach’s \( \alpha \) tends to underestimate the true reliability of test scores (Drewes, 2000; Nunnally & Bernstein, 1994; Osburn, 2000; Ten Berge & Hofstee, 1999) and shows a spurious decrease in its value when the number of categories in the rating scale is less than five (Lozano, García-Cueto, & Muñiz, 2008; Weng, 2004). An alternative procedure to Cronbach’s \( \alpha \) coefficient is ordinal \( \alpha \), based on the factor analysis of the polychoric correlation matrix (McDonald, 1999; Rupp, Koh, & Zumbo, 2003).

Construct validity has been traditionally explored using EFA, a statistical technique used to explore the possible underlying factor structure without imposing a preconceived structure on the outcome (Child, 1990). This technique is appropriate during test development, but confirmatory factor analysis (CFA) is a more powerful alternative that allows the researcher to postulate the relationship pattern a priori, using the knowledge of the theory or empirical research, and test the hypothesis statistically (Jöreskog, 1993).

The present research is aimed at providing updated evidence on the psychometric properties of the TPT in terms of reliability and construct validity. Reliability of the scales is explored using the ordinal approach, reliability of the dimensions is studied using the continuous factorial approach, and construct validity by means of CFA. We expected to replicate and to find better results with regards to reliability and construct validity than those reported in the technical manual. A satisfactory goodness-of-fit of the work personality model assessed by the TPT was also expected.

**Materials and Method**

**Sample**

Computerized archival data from the Internet Scoring System of TEA Ediciones (www.e-perfil.com) was used for this study. The TPT items and raw scores data were collected from 23,062 subjects with ages ranging from 18 to 75 (Mean = 30.10; SD = 7.62), assessed in Spain between 2005 and 2010. In the sample, 41.1% were female (\( n = 9,476 \)) and 58.9% were male (\( n = 13,586 \)).

**Instrument**

The TPT (Corral et al., 2002) is a 160-item self-report questionnaire to be answered using a 4-point Likert scale. According to the technical manual, TPT offers three main factor scores or dimensions (EE, Emotional Stability; AM, Mental Openness; and RP, Responsibility) and 15 facets of personality in the workplace: maladjustment (DAJ), anxiety (ANS), depression (DEP), stress tolerance (EST), self-concept (ACO), tolerance & flexibility (TOL), adaptation to changes (ADA), interest in other cultures (CUL), availability (DIS), social intelligence (ISO), social integration (INS), teamwork (EQUI), professional self-demanding (PRO), dynamism & activity (DIN), and tenacity & constancy (TES). It also includes a validity scale to assess sincerity (SIN). The TPT norms are based on a representative sample of 15,509 adults in the work context (23.3% males and 76.6% females).

**Data Analysis**

Polychoric matrix correlations of each scale and factor loadings of each item were obtained with MPLUS 5.2 software. Ordinal \( \alpha \)s were then calculated applying McDonald’s formula (1999, p. 217):

\[
\alpha = \frac{n}{n-1} \left[ \frac{m(\lambda)^2 - \overline{\lambda}^2}{m(\lambda)^2 + (u^2)} \right]
\]

where \( \overline{\lambda} \) is the arithmetical mean of factor loadings, \( \overline{\lambda}^2 \) is the arithmetical mean of the squared factor loadings, and \( u^2 \) is the arithmetical mean of the variable uniqueness.

Pearson matrix correlations of each second-order dimension and factor loadings of each scale were also obtained with MPLUS 5.2 software. Continuous factorial \( \alpha \) of dimensions were also calculated by means of Equation 1.

Differences between reliability coefficients were estimated with Feldt’s equality test with two independent samples (Feldt, 1980).

CFA with maximum likelihood (ML) estimation and fit statistics were done with AMOS 16.0. Two models were examined: (1) One single-factor control model with one first-order latent variable based on scales’ composites; and (2) One model with three correlated factors, with three first-order latent variables based on scales’ composites (related to the number of factors and scales loadings of the work personality model reported in the TPT technical manual).

**Procedure**

Data were obtained from the e-perfil system (www.e-perfil.com), a web scoring platform that collects and stores all the TPT administrations in Spain. Examiners use the paper and pencil version of the test and manually enter item responses into the system to obtain raw scores, standard scores, and the scoring profile. Data are stored and exported into Excel format and include administration date, gender, age, item responses, raw scores, and standard scores.
Table 1. Factor loadings for TPT scales: polychoric EFA, RMSR, ordinal α, and TPT’s technical manual Cronbach’s α

<table>
<thead>
<tr>
<th>Item</th>
<th>DAJ</th>
<th>ANS</th>
<th>DEP</th>
<th>EST</th>
<th>ACO</th>
<th>TOL</th>
<th>ADA</th>
<th>CUL</th>
<th>ISO</th>
<th>INS</th>
<th>DIS</th>
<th>EQU</th>
<th>PRO</th>
<th>DIN</th>
<th>TES</th>
<th>SIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.32</td>
<td>.35</td>
<td>.55</td>
<td>.42</td>
<td>.63</td>
<td>.45</td>
<td>.42</td>
<td>.26</td>
<td>.51</td>
<td>.31</td>
<td>.28</td>
<td>.26</td>
<td>.63</td>
<td>.48</td>
<td>.16</td>
<td>−.01</td>
</tr>
<tr>
<td>2</td>
<td>.65</td>
<td>.40</td>
<td>.50</td>
<td>.69</td>
<td>.43</td>
<td>.44</td>
<td>.36</td>
<td>.59</td>
<td>.60</td>
<td>.41</td>
<td>−.23</td>
<td>.33</td>
<td>.31</td>
<td>.35</td>
<td>.16</td>
<td>.22</td>
</tr>
<tr>
<td>3</td>
<td>.15</td>
<td>.12</td>
<td>.48</td>
<td>−.25</td>
<td>.23</td>
<td>.53</td>
<td>.02</td>
<td>.67</td>
<td>.21</td>
<td>.13</td>
<td>.68</td>
<td>.62</td>
<td>.54</td>
<td>.59</td>
<td>.64</td>
<td>.65</td>
</tr>
<tr>
<td>4</td>
<td>.25</td>
<td>.63</td>
<td>.33</td>
<td>.54</td>
<td>.63</td>
<td>.18</td>
<td>.64</td>
<td>.82</td>
<td>.48</td>
<td>.71</td>
<td>−.02</td>
<td>.31</td>
<td>.68</td>
<td>.41</td>
<td>.54</td>
<td>.11</td>
</tr>
<tr>
<td>5</td>
<td>.43</td>
<td>.61</td>
<td>.53</td>
<td>.46</td>
<td>.78</td>
<td>.36</td>
<td>.13</td>
<td>.31</td>
<td>−.06</td>
<td>.50</td>
<td>.42</td>
<td>.60</td>
<td>.37</td>
<td>−.04</td>
<td>.66</td>
<td>.12</td>
</tr>
<tr>
<td>6</td>
<td>.55</td>
<td>.64</td>
<td>.63</td>
<td>.24</td>
<td>.23</td>
<td>.51</td>
<td>.27</td>
<td>.49</td>
<td>.51</td>
<td>.34</td>
<td>.45</td>
<td>.42</td>
<td>.36</td>
<td>.47</td>
<td>.51</td>
<td>.55</td>
</tr>
<tr>
<td>7</td>
<td>.42</td>
<td>.60</td>
<td>.38</td>
<td>.44</td>
<td>.28</td>
<td>.38</td>
<td>−.06</td>
<td>.63</td>
<td>.62</td>
<td>.38</td>
<td>.67</td>
<td>.38</td>
<td>.69</td>
<td>.47</td>
<td>.41</td>
<td>.37</td>
</tr>
<tr>
<td>8</td>
<td>.57</td>
<td>.63</td>
<td>.64</td>
<td>.61</td>
<td>.63</td>
<td>.25</td>
<td>.13</td>
<td>.42</td>
<td>.64</td>
<td>.28</td>
<td>.39</td>
<td>.69</td>
<td>.59</td>
<td>.28</td>
<td>.68</td>
<td>.52</td>
</tr>
<tr>
<td>9</td>
<td>.27</td>
<td>.43</td>
<td>.60</td>
<td>.65</td>
<td>.17</td>
<td>.53</td>
<td>.76</td>
<td>.72</td>
<td>.59</td>
<td>.18</td>
<td>.33</td>
<td>.33</td>
<td>.67</td>
<td>−.07</td>
<td>.29</td>
<td>.17</td>
</tr>
<tr>
<td>10</td>
<td>.73</td>
<td>.15</td>
<td>.29</td>
<td>.64</td>
<td>.42</td>
<td>.09</td>
<td>.01</td>
<td>.41</td>
<td>.48</td>
<td>.56</td>
<td>.34</td>
<td>.33</td>
<td>.59</td>
<td>.29</td>
<td>.26</td>
<td>.31</td>
</tr>
<tr>
<td>Mean</td>
<td>.43</td>
<td>.45</td>
<td>.47</td>
<td>.44</td>
<td>.44</td>
<td>.37</td>
<td>.27</td>
<td>.53</td>
<td>.46</td>
<td>.38</td>
<td>.33</td>
<td>.43</td>
<td>.54</td>
<td>.32</td>
<td>.43</td>
<td>.30</td>
</tr>
<tr>
<td>RMSR</td>
<td>.050</td>
<td>.039</td>
<td>.081</td>
<td>.049</td>
<td>.087</td>
<td>.056</td>
<td>.115</td>
<td>.051</td>
<td>.088</td>
<td>.058</td>
<td>.064</td>
<td>.067</td>
<td>.053</td>
<td>.069</td>
<td>.086</td>
<td>.083</td>
</tr>
<tr>
<td>Ordinal α</td>
<td>.77</td>
<td>.77</td>
<td>.73</td>
<td>.72</td>
<td>.75</td>
<td>.64</td>
<td>.63</td>
<td>.86</td>
<td>.77</td>
<td>.73</td>
<td>.65</td>
<td>.76</td>
<td>.79</td>
<td>.58</td>
<td>.76</td>
<td>.55</td>
</tr>
<tr>
<td>Technical manual Cronbach’s α</td>
<td>.50</td>
<td>.62</td>
<td>.55</td>
<td>.64</td>
<td>.65</td>
<td>.61</td>
<td>.61</td>
<td>.78</td>
<td>.52</td>
<td>.69</td>
<td>.62</td>
<td>.62</td>
<td>.70</td>
<td>.49</td>
<td>.64</td>
<td>.51</td>
</tr>
</tbody>
</table>

Table 2. Factor loadings for TPT second order dimensions: EFA, RMSR, and α

<table>
<thead>
<tr>
<th>Scale</th>
<th>EE</th>
<th>AM</th>
<th>RP</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAJ</td>
<td>.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANS</td>
<td>.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEP</td>
<td>.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EST</td>
<td>.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACO</td>
<td>.71</td>
<td>.59</td>
<td></td>
</tr>
<tr>
<td>TOL</td>
<td>.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADA</td>
<td>.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUL</td>
<td>.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISO</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INS</td>
<td>.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIS</td>
<td>.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EQU</td>
<td>.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRO</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIN</td>
<td>.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TES</td>
<td>.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>.74</td>
<td>.59</td>
<td>.72</td>
</tr>
<tr>
<td>RMSR</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
</tr>
<tr>
<td>α</td>
<td>.89</td>
<td>.78</td>
<td>.87</td>
</tr>
</tbody>
</table>

Results

Consistency Analysis

Factor loadings for the 16 polychoric EFA (15 personality scales and 1 Sincerity scale), root mean square residuals (RMSR) and ordinal α values of the 10 items of the TPT scales are presented in Table 1; corresponding Cronbach’s αs reported in the TPT technical manual are also included. The loadings for the factors were achieved in separate analyses, where each item concentrates on one factor. Reliability in terms of ordinal α was calculated with Equation 1 and ranged from .55 to .86 (median = .74) and loadings ranged from −.25 to .76. Negative loadings can be observed in 8 of the 160 TPT items, corresponding to the following scales: EST, ADA, ISO, DIS, DIN, and SIN.

Factor loadings for the three continuous EFA, root mean square, and factorial α values are reported in Table 2. The loadings for these factors were also obtained in separate analyses, where each scale concentrates on its respective second-order factor. Reliability was also calculated with Equation 1 and values ranged from .78 to .89. All loading values were positive and ranged from .38 to .84.

Figure 1 shows the comparison between ordinal αs and Cronbach’s αs for the scales as described in the TPT technical manual. Ordinal coefficients of all scales were significantly higher than their respective original continuous values (p < .001), especially for the DAJ and DEP scales.

Structural Analysis

In order to test how well the theoretical work personality model of the TPT (reported in the technical manual) fitted the data collected, we compare that model to a more parsimonious model with just one second-order latent variable. The theoretical work personality model of the TPT comprises 3 second-order factors with the following structure: EE includes DAJ, ANS, DEP, EST, and ACO; TOL, ADA, CUL, ISO, INS, and ACO constitute the AM factor; and DIN, PRO, EQU, TES, and DIS constitute the RP factor. Two CFAs with ML estimation were used. According to Schweizer’s recommendations (Schweizer, 2010), the following goodness-of-fit indexes were calculated for the single factor model: $\chi^2 = 3872.716$, df = 90, normed $\chi^2 = 43.03$, root mean square error of approximation (RMSEA) = .10 (.099 to .104), adjusted goodness-of-fit index (AGFI) = .83, comparative fit index (CFI) = .87, and standardized
root mean square residual (SRMR) = .05; the three-factor model had the following goodness-of-fit indexes: \( \chi^2 = 2735.849, df = 86, \) normed \( \chi^2 = 31.81, \) RMSEA = .08 (.084 to .090), AGFI = .89, CFI = .92, and SRMR = .05. These findings indicate a better fit of the theoretical model to empirical data than the more parsimonious model. Figure 2 presents the standardized regression weights and the structure of the theoretical model.

Discussion

The aim of this study was to study the psychometric properties of the TPT and to provide updated, more accurate, and representative information about its properties, that is, its reliability and construct validity.

The first hypothesis that predicted an increase in the reliability of the TPT scales was confirmed according to the data analyzed (\( p < .001 \)). Part of this increase is a result of the ordinal procedure used for its calculation, but it is widely argued that this is a more accurate reliability procedure, taking into account the Likert scale, than the equivalent procedure used and reported in the technical manual (Drewes, 2000; Elosúa & Zumbo, 2008; Lozano et al., 2008; Nunnally & Bernstein, 1994; Osburn, 2000; Ten Berge & Hofstee, 1999; Weng, 2004). These reliability results were expected; the TPT would hardly be one of the most commonly used tests in Spain (Muñiz & Fernandez-Hermida, 2010) if it lacked enough reliability to guarantee the required assessment quality. However, 5 of the 16 scales had reliabilities that could be improved according to the ordinal \( \alpha \) and loadings achieved. Two of them seems to be improvable with an extensive revision of the negative-loading items. That is the case for DIS and DIN. Moreover, TOL, ADA, and SIN could be indicating a comprehensive review of the overall item content.

Results related to the reliability of second-order dimensions were satisfactory; therefore, the TPT is a measuring tool with an acceptable (and improvable) accuracy of its scales and a good reliability of their second-order dimensions.

The second hypothesis tested was also confirmed by the empirical data. The TPT’s psychological structure, consisting of three underlying second-order dimensions (Emotional Stability, Mental Openness and Responsibility), has been confirmed by the empirical data. According to the CFA results, the psychological content of the three dimensions can be better described. Emotional Stability is defined primarily by low levels of Anxiety, Depression, and Adjustment in the workplace, and as a good ability to manage stress and a good self-concept. This dimension has been shown to be an important trait of the work personality in other relevant tools developed in Spain (compeTEA, Arribas & Pereña, 2009) and Europe (BFQ; Caprara, Barbaranelli, & Borgogni, 1993; BIP; Hossiep & Paschen, 2008).

Mental Openness is a particularly interesting dimension because it is one of the Big Five model traits traditionally ascribed to the normal personality (McCrae & Costa, 1987, 1997) and seems to be important for the work personality as well. To be an open-minded person in the TPT parameters means to be tolerant and flexible, adaptive to changes, interested in other cultures, and with a high level of social intelligence and social integration in the company.

Responsibility is the factor most clearly related to performance at work (Salgado, 2003). Also included in other classic questionnaires such as NEO PI-R (Costa & McCrae, 1992) or SOSIA (Gordon, ECPA, 1990; Gordon, ECPA, & TEA Ediciones, 1997), Responsibility consists of a set of characteristics that define a “good worker”: dynamism and activity, professional self-demanding, availability, tenacity, and constancy and teamwork.

A further interesting result was obtained. Most of the items of the TPT present high or moderate loadings on their respective scale, which influences the value of the reliability coefficient. However, eight of the items had a negative
loading on the EST, ADA, DIS, ISO, DIN, and SIN scales. It would be interesting to carry out future studies to examine and review the content of these items to increase the reliability of these scales. Additional study of the content of the scales that had lower loadings and reliabilities would also be interesting.

Despite the above encouraging results, there are some limitations to this study. First, because of the data collection system, the main purpose of the assessment was recruiting employees, but no data was available regarding the specific vacancy for which subjects were evaluated. Hence, the effect of this variable on the psychometric properties is unknown. Furthermore, it would have been interesting to have the ordinal data reliability from the technical manual. Finally, the TPT’s theoretical model appropriately fits the empirical data for Spain. Future research could explore the usefulness of this model for other European countries.

References


Méndez, F., Hidalgo, M., & Inglés, C. (2002). The Matson Eval...


David Arribas-Águila

Research and Development Department
TEA Ediciones
c/Fray Bernardino de Sahagún 24
28036 Madrid
Spain
E-mail david.arribas@teaediciones.com