NORMATIVE DATA, RELIABILITY, AND VALIDITY OF THE NEO PI-R PERSONALITY DISORDER SCALES

Ana Sanz-García, María Paz García-Vera, and Jesús Sanz
Complutense University of Madrid (Spain)

Abstract

There is increasing interest in understanding personality disorders (PDs) from the five-factor model. Miller et al. (2005) and Costa and McCrae (2005) proposed two sets of scales based on the NEO Personality Inventory-Revised (NEO PI-R) facets to assess DSM-5 PDs. There are Spanish norms for the scales of Miller et al. (2005) based on personnel selection samples, but they are not appropriate for contexts with low social desirability. Normative, reliability, and convergent/discriminant validity data are presented for both sets of scales with volunteers from the general Spanish population (*N*= 682). The internal consistency and convergent/discriminant validity indices were excellent or good for all scales, especially for those of Miller et al. (2005). The differences between the sample of volunteers and that of personnel selection (*d*= 0.61) and between males and females (*d*= 0.34-0.38) justify the development of norms for the two sets of PD scales for situations of low social desirability and separate for males and females. Their usefulness in different contexts is discussed.

KEY WORDS: *personality disorders, five-factor model, NEO PI-R, normative data, reliability, validity*.

Resumen

Hay un creciente interés por entender los trastornos de personalidad (TTPP) desde el modelo de los cinco factores. Miller *et al.* (2005) y Costa y McCrae (2005) propusieron dos conjuntos de escalas basadas en las facetas del "Inventario de personalidad NEO-revisado" (NEO PI-R) para evaluar los TTPP del DSM-5. Existen baremos españoles para las escalas de Miller *et al.* (2005) a partir de muestras de selección de personal, pero no son apropiados en contextos con deseabilidad social baja. Se presentan datos normativos, de fiabilidad y validez convergente/ discriminante para ambos conjuntos de escalas con voluntarios de la población general española (*N*= 682). Los índices de consistencia interna y validez convergente/discriminante fueron excelentes o buenos para todas las escalas, especialmente para las de Miller *et al.* (2005). Las diferencias entre la muestra de voluntarios y de selección de personal (*d*= 0,61) y entre varones y mujeres (*d*= 0,34-

---

0,38) justifican el desarrollo de baremos para los dos conjuntos de escalas de TTPP para situaciones de deseabilidad social baja y separados para varones y mujeres. Se discute su utilidad en diferentes contextos.

Palabras clave: *trastornos de la personalidad, modelo de los cinco factores, NEO PI-R, baremo, fiabilidad, validez.*

Introduction

Currently, the five-factor model (FFM) of personality or Big Five model is the most valid, consensual and widely used taxonomy of normal personality traits because, for example, it has been replicated in different countries and languages and with different instruments and populations (John, 2021; McCrae, 2020; Sanz, 2018). The Revised NEO Personality Inventory or NEO PI-R (Costa & McCrae, 1992) has become the standard for FFM assessment (Costa & McCrae, 2009; Sanz, 2018; Sanz et al., 1999; Sanz-García et al., 2023). In fact, the thousands of published studies that have used the NEO PI-R are one of the most important sources of evidence supporting the validity of the FFM (Sanz, 2018). The NEO PI-R allows us to obtain measures of each of the five dimensions of the FFM —neuroticism, extraversion, openness to experience, agreeableness and conscientiousness— but also of the 30 facets or specific traits —six facets for each dimension— that make up these dimensions according to the FFM of Costa and McCrae (2009; McCrae & Costa, 2003).

One of the virtues of the FFM is its ability to compare, contrast, and integrate personality constructs that measure seemingly very disparate personality assessment instruments (Costa & McCrae, 2009; McCrae & Costa, 2003; O'Connor, 2017; Sanz, 2018; Sanz et al., 2008). In this sense, extensive scientific literature has also shown that the FFM is capable of integrating the characteristics and symptoms of PDs included in the latest editions of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV and DSM-5) of the American Psychiatric Association (1994/1995, 2013/2014) (Widiger et al., 2017). In fact, each PD of both the DSM-IV and DSM-5, as both editions share the same categories and diagnostic criteria for such disorders, has a distinctive profile according to the FFM and, in particular, according to the facets of the NEO PI-R. These profiles have been obtained both from the consensus of groups of expert researchers (Lynam & Widiger, 2001) or clinicians (Samuel & Widiger, 2004) and from meta-analyses of empirical studies that relate measures of the FFM, most of which were provided by the NEO PI-R, with measures of PDs from the DSM-IV (Samuel & Widiger, 2008). In addition, both profiles, those based on expert consensus and the empirically based ones, largely coincide with correlations ranging from .60 for dependent PD to .92 for obsessive-compulsive PD (Samuel & Widiger, 2008).

Based on the prototypical profiles agreed upon by the experts of the study of Lynam and Widiger (2001), Miller et al. (2005) created a set of 10 scales to assess the different PDs based on the NEO PI-R facets. In the same vein, based on the

theoretical translation of the DSM-IV diagnostic criteria for PDs in terms of the NEO PI-R facets that Widiger et al. (2002) made of the diagnostic criteria of the DSM-IV for PDs, Costa and McCrae (2005) also developed a set of 10 scales based on these facets to assess these disorders. In both sets, the scales add the direct or raw scores of the facets positively related to the PD in question and subtract the scores (or add the inverted scores) of the negatively related facets. Several studies have obtained evidence of convergent, discriminant, and diagnostic validity for the scores in both the first set (Miller et al., 2005, 2008) and the second set of scales (Costa & McCrae, 2005).

In Spain, Colodro et al. (2018) presented evidence of internal consistency reliability, and convergent and discriminant validity of the set of scales of Miller et al. (2005) and developed norms for them based on the official Spanish adaptation of the NEO PI-R (Costa & McCrae, 2008). This adaptation was carried out with samples evaluated during personnel selection processes, so, as Sanz and García-Vera (2009) argue, its norms seem appropriate for contexts in which it is assumed that people are motivated to present a favorable image of themselves; that is, contexts in which a significant bias of social desirability is suspected. However, they do not seem suitable for situations where social desirability is presumed to be low or non-existent, such as in research conducted with volunteers or in applications conducted in many clinical or educational settings. In fact, the results of Sanz and García-Vera (2009) show that the norms of the official Spanish adaptation of the NEO PI-R (Costa & McCrae, 2008) differ markedly (0.56 – 0.51 standard deviation units on average) from the mean scores obtained by adult volunteers from the general population, so the use of those norms can significantly distort the interpretation of the scores of an adult who voluntarily completes the NEO PI-R, presenting lower levels of extraversion, openness and agreeableness and, especially, higher levels of neuroticism and lower levels of conscientiousness, than they actually have compared to the Spanish reference population.

Therefore, when developing the norms of the PD scales of Miller et al. (2005) based on the norms of the official Spanish adaptation of the NEO PI-R, the norms of Colodro et al. (2018) for these PD scales would be appropriate for situations in which social desirability is suspected to be high (e.g., promotion or job selection processes, forensic evaluations in child custody disputes), but they would not be suitable for situations where low or no social desirability is presumed, as they could lead to a spurious rise in the standardized scores of the people assessed and, consequently, an overestimation of the presence of PD. In fact, in a sample of adults who had voluntarily participated in research comprising mostly university students (69.5%), Colodro et al. (2018) obtained, from their norms, PD prevalence rates that, except for obsessive-compulsive PD (0%), ranged from 9% (schizoid PD) to 38% (antisocial PD), so 70% of these adults presented some PD. These rates appear to significantly overestimate the prevalence of PDs in the general adult population, as the rates obtained in epidemiological studies with samples from this population are much lower. For example, the mean rates Winsper et al. (2020) found in their meta-

analysis of 46 epidemiological studies ranged from only 0.6% (histrionic PD) to 3.2% (obsessive-compulsive PD), so only 7.8% of adults in the general population presented some PD. Similarly, the results of the meta-analysis of Volkert et al. (2018), obtained with only 10 studies but all of them carried out with samples of the general adult population of Western countries and, therefore, more homogeneous and similar to the general Spanish population, also indicate much lower prevalence rates of PDs than those found by Colodro et al. (2018) because Volkert et al. (2018) found average prevalence rates ranging only from 0.8% (histrionic and dependent PDs) to 4.3% (obsessive-compulsive PD), so only 12.2% of adults presented some PD.

Consequently, the present study had three related objectives. The first is to present normative and reliability data for the PD scales of the NEO PI-R of Miller et al. (2005), obtained with a heterogeneous sample of adult volunteers from the general Spanish population, which allows the use of these scales in research contexts with volunteers or in situations where social desirability is presumed to be low or non-existent, such as, for example, some clinical or educational contexts. The second objective is to present normative and reliability data for the PD scales of the NEO PI-R of Costa and McCrae (2005), which, to our knowledge, have not yet been investigated in Spain. The third objective is to analyze the convergent and discriminant validity of the scores of both sets of PD scales.

Method

Participants

This study involved 682 adults (390 females and 292 males) aged between 18 and 84 years (*M*= 41.8, *SD*= 14.8) whose responses to the NEO PI-R were used in a previous study on the evaluation of this instrument in the general Spanish population (Sanz & García-Vera, 2009). These people were recruited in 2002-2004 using the snowball technique by university students of Psychology who invited their family and friends to voluntarily participate in a study on personality and hypertension (*n*= 358) or another on personality assessment (*n*= 325) although the university students of Psychology themselves were not part of the samples of these two studies nor, consequently, of the present study. Although a sample obtained this way is not random, its profile concerning sex and age was very similar to that of the Spanish population in 2004 (see table 1 of Sanz & García-Vera, 2009). More information on the sociodemographic characteristics of the sample can be found in Sanz and García-Vera (2009), where it can be seen that the sample was also heterogeneous in terms of educational level, marital status and profession (e.g., 22.1% of the participants had primary education as their highest level of education, 30.2% had secondary education, and 45.3% had university studies; 55.9% of the participants were married or living with a stable partner, 35% were single, 6% were separated or divorced, and 2.3% were widowed).

Instruments

Revised NEO Personality Inventory (NEO PI-R; Costa & McCrae, 1992). The NEO PI-R is a 240-item self-reporting instrument rated on 5-point Likert-type scales, ranging from 0 to 4, designed to evaluate personality based on the FFM. The NEO PI-R has five basic scales, each composed of 48 items, which correspond to the basic dimensions of the Big Five: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. The NEO PI-R also has 30 specific scales of 8 items each (six for each basic scale) that aim to measure the facets or specific personality factors that, according to Costa and McCrae (1992), make up the Big Five: Anxiety, Angry Hostility, Depression, Self-Consciousness, Impulsiveness, Vulnerability, Warmth, Gregariousness, Assertiveness, Activity, Excitement-Seeking, Positive Emotions, Fantasy, Aesthetics, Feelings, Actions, Ideas, Values, Trust, Straightforwardness, Altruism, Compliance, Modesty, Tender-Mindedness, Competence, Order, Dutifulness, Achievement Striving, Self-Discipline, and Deliberation. In this study, the official Spanish adaptation of the NEO PI-R was used (Costa & McCrae, 1999). In the present sample of the general Spanish population, the scores in the NEO PI-R dimensions obtained alpha reliability coefficients between .86 and .90, whereas the facet scores reached alpha reliability coefficients between .49 and .81, with a median of .67.

Procedure

Participants who collaborated in the personality and hypertension research completed the NEO PI-R as part of a more comprehensive assessment in which they had to fill out other personality questionnaires, with the NEO PI-R as the first. Participants who collaborated in the personality assessment research only completed the NEO PI-R. In both cases, the NEO PI-R was applied individually and on paper by the psychology student who had invited the participant to collaborate in one of these two investigations as part of his practical classes or a seminar. The training and supervision of the students in the administration of the NEO PI-R was carried out by the last two authors of this study during those practical classes or seminars.

Data analysis

To calculate the participants' scores on the two sets of PD scales based on the NEO PI-R facets, the formulas in table 14.2 of Costa & McCrae (2005) were used, on the one hand, and, on the other hand, like Colodro et al. (2018), the formulas in the Appendix of Miller et al. (2005) and, for the dependent PD, the updated formula of Miller and Lynam (2008).

Central tendency (mean), dispersion (standard deviation) statistics, and normality (kurtosis and skewness) of the distribution of scores for the two sets of PD scales were calculated. The internal consistency reliability of these scores was

analysed using Cronbach's alpha coefficient, and their convergent and discriminant validity with Pearson's correlations between the corresponding scales of both sets (convergent correlations) and the mean of the correlations between each PD scale and the scales of the other nine PDs in both sets; that is, the mean of 18 discriminant correlations. These means were calculated by applying Fisher's transformation to the correlations, averaging the resulting Fisher z-values, and converting the mean of these Fisher z-values into a correlation. After applying Fisher's transformation to the correlations involved in the difference (q= Fisher's z_1 – Fisher's z_2), the differences between the two types of correlations, the convergent correlations, and the means of the discriminants were interpreted according to Cohen's (1988) standards for the size of the differences between correlations: $0.10 \leq q < 0.30$, $0.30 \leq q < 0.50$, and $q \geq 0.50$ were considered small, medium, and large differences, respectively.

In the case of the scales of Miller et al. (2005; Miller & Lynam, 2008), using Student's t-tests, the mean scores obtained in the present sample of participants were compared to those obtained by Colodro et al. (2018) based on the staff selection sample of the official Spanish adaptation of the NEO PI-R (Costa & McCrae, 2008), and the magnitudes of the differences were estimated using Cohen's d effect size statistic.

As Sanz and García-Vera (2009) found statistically significant differences between males and females in most of the NEO-PI-R facets and, in some of them, such differences were close to average effect sizes, Student's t-tests were conducted to analyze the sex differences on the PD scales and, if they were statistically significant, differentiated norms were calculated for males and females for the two sets of PD scales.

Results

Distribution and internal consistency of NEO PI-R personality disorder scale scores

Table 1 presents the descriptive statistics of the scores of the two sets of PD scales, as well as their internal consistency indices. All the scales had kurtosis and skewness values within the range considered indicative of a normal distribution of scores (± 1), except for the Paranoid PD scale of Costa and McCrae (2005) and the two scales of Narcissistic PDs, which obtained kurtosis values greater than 1, although less than 1.5.

According to the standards of Hernández et al. (2016), all the scales of Miller et al. (2005) had excellent internal consistency indices ($\alpha \geq$ .85), except for the Schizotypal PD, which presented a good index (.80 $\leq \alpha <$ .85). Five of the ten scales of Costa and McCrae (2005) had excellent indices, another four had good indices and the remaining scale, Narcissistic PD, had an adequate index (.70 $\leq \alpha <$ .80).

Table 1

Mean (*M*), standard deviation (*SD*), kurtosis and skewness indices and internal consistency (α coefficient) of the scores of the personality disorder (PD) scales of the NEO PI-R proposed by Miller et al. (2005) and Costa and McCrae (2005)

PD scales	*M*	*DT*	Skewness	Kurtosis	α
Miller et al. (2005)					
Paranoid	125.23	24.76	0.50	0.89	.88
Schizoid	115.10	24.15	0.09	-0.08	.89
Schizotypal	101.49	17.65	0.14	0.41	.81
Antisocial	225.23	32.47	0.21	0.46	.87
Borderline	136.62	23.34	0.16	0.02	.86
Histrionic	200.10	28.98	0.00	-0.00	.88
Narcissistic	166.64	27.18	0.39	1.34	.86
Avoidant	159.69	25.74	0.04	-0.11	.87
Dependent*	84.09	22.06	0.18	-0.02	.91
Obsessive-compulsive	227.52	29.25	-0.12	-0.14	.87
Costa y McCrae (2005)					
Paranoid	51.94	13.53	0.50	1.01	.83
Schizoid	49.36	14.85	0.24	0.05	.87
Schizotypal	132.69	19.54	0.18	0.53	.80
Antisocial	107.29	23.55	0.21	0.45	.87
Borderline	112.05	23.31	0.14	0.10	.88
Histrionic	160.00	21.47	-0.03	0.26	.83
Narcissistic	96.87	13.99	0.38	1.46	.71
Avoidant	110.35	23.07	0.16	-0.00	.89
Dependent	163.31	21.62	-0.11	0.75	.85
Obsessive-compulsive	123.26	17.26	0.07	0.13	.82

Note: *Updated scale of Miller and Lynam (2008).

Convergent and discriminant validity of the scores of the NEO PI-R personality disorders scales

Table 2 presents the convergent and discriminant validity coefficients of the scores of the two sets of PD scales. According to the standards of Hernández et al. (2016), all the PD scales showed excellent convergent validity indices ($r \geq .70$), except for the Narcissistic and Obsessive-Compulsive PD scales, which showed good indices ($.60 \leq r < .70$), and the Dependent PD scales, which showed an adequate index ($.50 \leq r < .60$). For all PD scales, these convergent correlations were greater than the means of the 18 discriminant correlations, and, moreover, these differences were all large ($q \geq 0.50$) except for two, the Dependent PD scale of Miller et al. (2005) and the Narcissistic PD scale of Costa and McCrae (2005), which were medium ($0.3 \leq q < 0.5$). Moreover, for most scales, none of their 18 discriminant correlations was greater than the corresponding convergent correlation with a difference (q) greater than 0.10, except for the Narcissistic and Dependent PD scales of Miller et al. (2005),

which, nonetheless, presented only three of the 18 discriminant correlations greater than their corresponding convergent correlation with a difference higher than 0.10.

Table 2
Convergent and discriminant validity of the scores of the personality disorder (PD) scales of the NEO PI-R

PD scales	Convergent correlation (r_c)	Discriminant correlations (mean r_d)	Effect size (q)	No. of r_d > r_c
Miller et al. (2005)				
Paranoid	.80	.30	**0.789**	0
Schizoid	.91	.07	**1.458**	0
Schizotypal	.88	.30	**1.066**	0
Antisocial	.87	.10	**1.233**	0
Borderline	.89	.29	**1.123**	0
Histrionic	.78	-.24	**1.290**	0
Narcissistic	.66	.18	**0.611**	3
Avoidant	.90	.16	**1.311**	0
Dependent*	.59	.33	0.335	3
Obsessive-compulsive	.61	-.19	**0.901**	0
Costa y McCrae (2005)				
Paranoid	.80	.30	**0.789**	0
Schizoid	.91	.12	**1.407**	0
Schizotypal	.88	.32	**1.044**	0
Antisocial	.87	.26	**1.067**	0
Borderline	.89	.38	**1.022**	0
Histrionic	.78	-.15	**1.195**	0
Narcissistic	.66	.30	0.483	0
Avoidant	.90	.27	**1.195**	0
Dependent	.59	-.10	**0.778**	0
Obsessive-compulsive	.61	-.16	**0.870**	0

Notes: * Updated scale of Miller and Lynam (2008). The differences between the convergent correlation and the mean of the 18 discriminant correlations that are medium (q> .30) or large (q> .50) are shown in bold. To consider a discriminant correlation to be greater than the convergent correlation, its difference (q) must be greater than 0.10.

Differences in the PD scales of the NEO PI-R between adult volunteers and those assessed in personnel selection processes

Table 3 shows the standardized differences (Cohen's *d*) in the PD scales of the NEO PI-R of Miller et al. (2005, 2008) between the means of the sample of Spanish volunteers in this study and the normative means obtained by Colodro et al. (2018) from the personnel selection norms of the official Spanish adaptation of the NEO PI-R (Costa & McCrae, 2008). The results of the *t*-tests revealed that all differences were statistically significant at *p*< .0001. The sizes of the differences ranged from -0.20 (Histrionic PD) to 1.00 (Paranoid PD), large (*d*≥ 0.80) for three of the 10 scales,

medium (0.50≤ *d*< 0.80) for another three, and small (0.30≤ *d*< 0.50) for three of the remaining four. Hence, the average of the differences, in absolute value, for all scales represented a medium effect size (*d*= 0.61) and, as might be expected, the differences indicated a more socially favorable image for the recruitment sample; that is, a higher score on the PDs scales in the sample of Spanish volunteers except for the Histrionic and Obsessive-Compulsive PD scales.

However, when the means of the sample of Spanish volunteers in this study were compared with those obtained in other countries with volunteers from the general population, in particular the United States, France, and Belgium-Netherlands (Table 3), none of the 30 possible differences were large, only one was medium-sized, and only seven were small.

Table 3

Standardized difference (Cohen's *d*) in the personality disorder scales of the NEO PI-R proposed by Miller et al. (2005) between the means of this study and other studies

Personality disorder scales of Miller et al. (2005)	Spain – general population compared to:			
	Spain – Personnel selection	USA – General population	France – General population	Belgium-Netherlands – General population
Paranoid	**1.00**	-0.28	-0.07	0.04
Schizoid	**0.68**	-0.23	-0.32	0.07
Schizotypal	**0.83**	-0.27	0.09	-0.19
Antisocial	0.49	0.19	0.36	0.07
Borderline	**0.93**	-0.33	0.31	-0.21
Histrionic	-0.20	0.13	0.31	-0.02
Narcissistic	0.41	0.02	0.13	0.01
Avoidant	**0.74**	-0.47	-0.16	-0.14
Dependent*	0.31	-0.32	-0.02	-0.21
Obsessive-compulsive	**-0.55**	-0.26	**-0.50**	-0.10

Notes: *Updated scale of Miller and Lynam (2008). Standardized differences in medium-sized (0.50≤ *d*< 0.80) and large-sized means (*d*≥ 0.80) are shown in bold.

Sex differences in the PD scales of the NEO PI-R and norms for Spanish adult volunteers

Table 4 presents the means and standard deviations necessary to transform the raw scores into standardized scores. In addition, as can be seen in Table 4, the *t*-tests of mean differences between sexes revealed statistically significant differences between males and females for all the PD scales of the NEO PI-R except for the Obsessive-Compulsive PD scale of Miller et al. (2005) and the Schizotypal PD scale of Costa and McCrae (2005). For half of them (10 out of 20), such differences reached or exceeded medium effect sizes (*d*≥ 0.50) or approached medium sizes (*d* > 0.40).

Therefore, Appendices 1-4 show percentiles of the NEO PI-R PD scales for adult volunteers, differentiated for males and females.

Table 4
Mean and standard deviation of the scores of the personality disorder scales of the NEO PI-R by sex and standardized mean sex difference (Cohen's *d*)

Personality disorder scales	Males		Females		Cohen's
	M	*SD*	*M*	*SD*	*d*
Miller et al. (2005)					
Paranoid	133.01	24.81	119.41	23.10	0.57***
Schizoid	117.89	24.51	113.01	23.70	0.20**
Schizotypal	103.38	17.09	100.08	17.95	0.19*
Antisocial	234.25	31.62	218.48	31.49	0.50***
Borderline	130.98	21.82	140.86	23.57	-0.43***
Histrionic	197.43	29.44	202.11	28.50	-0.16*
Narcissistic	176.06	26.06	159.60	25.86	0.63***
Avoidant	155.35	25.48	162.96	25.49	-0.30***
Dependent*	79.08	20.74	87.86	22.30	-0.41***
Obsessive-compulsive	225.45	29.91	229.08	28.69	-0.12
Costa y McCrae (2005)					
Paranoid	54.27	13.93	50.20	12.96	0.30***
Schizoid	52.97	14.79	46.66	14.33	0.43***
Schizotypal	134.09	18.60	131.64	20.17	0.12
Antisocial	112.82	22.67	103.15	23.38	0.42***
Borderline	107.96	21.91	115.12	23.87	-0.31***
Histrionic	154.99	21.43	163.76	20.75	-0.42***
Narcissistic	99.16	13.85	95.16	13.87	0.29***
Avoidant	104.41	21.78	114.80	23.03	-0.46***
Dependent	153.89	19.77	170.38	20.24	-0.82***
Obsessive-compulsive	125.21	17.87	121.81	16.67	0.20*

Note: [a]Updated scale of Miller and Lynam (2008). ***Significant mean difference with *p*< .001. **Significant mean difference with *p*< .01. *Significant mean difference with *p*< .05.

Cut-off scores and prevalence of PDs in Spanish adult volunteers

To identify people who may have a PD, Miller et al. (2008) proposed a standardized T-score of 65 as the cut-off score for the PD scales of the NEO PI-R of Miller et al. (2005). For their scales, Costa and McCrae (2005) proposed using as a cut-off score the score that, in a normative sample, obtains or exceeds the same percentage of people that reproduces the percentage of people who actually have the PD in question in the general population, for which they considered the prevalence rates of the PDs provided by the DSM-IV (American Psychiatric Association, 1994). This last proposal is similar to the base rate (BR) scores used by the different versions of the Millon Multiaxial Clinical Inventory or MCMI (Sanz, 2007). In addition, Costa and McCrae (2005) proposed using the median of a

normative sample to identify people who, by scoring below that cut-off point, are unlikely to suffer from a PD.

Table 5
Cut-off scores and prevalence of personality disorders in the general population sample

Personality disorder	Median	Prevalence of the meta-analysis of Winsper et al. (2020)	BR	Prevalence with BR	T 65	Prevalence with T≥ 65
Miller et al. (2005)						
Paranoid	124	2.3%	180	2.2 %	162	6.3%
Schizoid	115	1.1%	174	1.3 %	151	6.9%
Schizotypal	101	0.8%	148	0.9 %	128	7.2%
Antisocial	224	1.4%	298	1.3 %	274	6.3%
Borderline	136	1.8%	189	1.8 %	172	6.7%
Histrionic	198	0.6%	271	0.6 %	244	6.6%
Narcissistic	166	1.9%	234	1.9 %	207	6.7%
Avoidant	160	2.7%	209	2.5 %	198	7%
Dependent*	83	0.8%	141	0.9 %	117	6.9%
Obsess.-comp.	229	3.2%	282	3.1 %	271	6.7%
Any disorder		7.8%		11.3%		34.9%
Costa and McCrae (2005)						
Paranoid	51	2.3%	83	2.2 %	72	8.1%
Schizoid	48	1.1%	89	1.0 %	72	6.5%
Schizotypal	132	0.8%	181	0.9 %	162	7.6%
Antisocial	106	1.4%	167	1.2 %	143	7.3%
Borderline	111	1.8%	163	1.8 %	147	7.3%
Histrionic	160	0.6%	214	0.6 %	192	7.0%
Narcissistic	96	1.9%	131	1.9 %	118	6.0%
Avoidant	110	2.7%	155	2.8 %	145	7.6%
Dependent	163	0.8%	216	0.9 %	196	6.5%
Obsess.-comp.	123	3.2%	156	3.1 %	149	7.5%
Any disorder		7.8%		11.1%		36.1%

Notes: *Updated scale of Miller and Lynam (2008). Obsess.-comp.= Obsessive-compulsive; Median= median of the normative data of the present study and cut-off score below which the presence of PD is unlikely. BR= Base rate cut-off score based on the prevalence of the disorder in the general population according to the meta-analysis of Winsper et al. (2020) and above which the presence of the PD is likely. Prevalence with BR= prevalence of the disorder in the sample of the present study using the BR score as the cut-off score. T 65= T-score of 65 based on the normative data of this study. Prevalence with T≥ 65= prevalence of the disorder in the sample of the present study using the T-score 65 as the cut-off score.

Following these proposals, Table 5 presents, for both sets of scales and calculated from the sample of the present study, the T-scores of 65, the medians, and the BR scores, although to calculate the latter, we considered the prevalence rates of the PDs of the meta-analysis of Winsper et al. (2020), much more up-to-

date than those of the DSM-IV. Table 5 also shows the prevalence rates of PDs obtained in the sample of this study using the T-scores of 65 and the BR. As might be expected given the normal distribution of the scores on the two sets of scales, the prevalence rates were higher for the T of 65 and very similar for both sets of scales, ranging from 6.3% (Paranoid PD and Antisocial PD) to 7.2% (Schizotypal PD) for the scales of Miller et al. (2005) and between 6% (Narcissistic PD) and 8.1% (Paranoid PD) for the scales of Costa and McCrae (2005), such that between 34.9% (scales of Miller et al., 2005) and 36.1% (scales of Costa & McCrae, 2005) of the participants had some PD. In contrast, the prevalence rates using the BR as a cut-off score ranged from 0.6% (Histrionic PD) to 3.1% (Obsessive-Compulsive PD) for both sets of scales, with about 11.1 – 11.3% of participants presenting a PD.

Discussion

As demonstrated by the incorporation into the DSM-5 of a dimensional model of PDs as an alternative model to the classical categorical perspective (American Psychiatric Association, 2013/2014), there has been growing interest in understanding these disorders from the perspective of the dimensions and traits of normal personality, especially from the perspective of the FFM (Watson & Clark, 2020; Widiger & Costa, 2013). Miller et al. (2005) and Costa and McCrae (2005) proposed two sets of 10 scales for the assessment of PDs based on the facets measured by the NEO PI-R. This inventory is currently the standard instrument for the assessment of the personality dimensions and traits of the five-factor model or the Big Five. In this context, the main objectives of this study were to present normative, reliability, and convergent and discriminant validity data for these two sets of scales with a heterogeneous sample of adult volunteers from the general Spanish population.

The results of the present study suggest that the scores of both sets of scales have reliability indices of internal consistency (α) that, according to the standards of Hernández et al. (2016), can be considered excellent or good in all of them, especially for the scales of Miller et al. (2005) compared to those of Costa and McCrae (2005) (median α coefficients= .87 vs .84, respectively). The sole exception is the Narcissistic PD scale of Costa & McCrae (2005), which, nevertheless, presents an adequate index. These indices are slightly higher than those obtained by Colodro et al. (2018) for the scales of Miller et al. (2005) both in a non-clinical sample composed mostly of university students and in a clinical sample of outpatients with psychological disorders. They obtained α coefficients ranging from .52 to .83, with a median of .70, for the non-clinical sample and between .58 and .84, with a median of .71, for the clinical sample. In the sample of the present study, for the scales of Miller et al. (2005), α coefficients ranging from .81 to .91 were obtained, with a median of .87. In any case, the results of the present study are consistent with those of the study of Colodro et al. (2018) in corroborating the internal consistency reliability of the scores of the PD scales of Miller et al. (2005) in samples of Spanish

adults and, in addition, they provide the novelty of showing the internal consistency reliability of the scores of the PDs scales of Costa and McCrae (2005).

The results of the present study also suggest that the scores of most of the scales of Miller et al. (2005) and of Costa and McCrae (2005) have excellent indices of convergent validity ($r \geq$ .70), which, in addition, are greater, with a large difference, than the means of the discriminant correlations. In general, the scales of Narcissistic and Dependent PDs are the exceptions to this pattern of convergent-discriminant correlations. However, these two exceptions show good or adequate convergent validity indices that are also higher than the means of the discriminant correlations, although in some cases, the difference was of medium size rather than large. For the scales of Miller et al. (2005), these results are consistent with the indices of convergent and discriminant validity obtained by Colodro et al. (2018) in two samples of Spanish adults, one clinical and the other non-clinical, and with those obtained by Miller et al. (2005) with a clinical sample. However, in both studies, the indices of convergent validity were lower because the convergence criterion was a measure of PDs provided by an instrument other than the NEO PI-R. For the scales of Costa and McCrae (2005), the results of the present study are also consistent with the convergent validity indices obtained by Costa and McCrae (2005) with a clinical sample, although, again, the latter indices were lower because the convergence criterion was also a measure of PDs provided by an instrument other than the NEO PI-R.

In the sample of Spanish adults from the general population that participated in the present study, statistically significant sex differences were found for practically all the PD scales of Miller et al. (2005) and Costa and McCrae (2005), on average, of medium size for both sets of scales ($d =$ 0.38 and 0.34, respectively), whereas Colodro et al. (2018) reported finding no relevant sex differences ($d =$ 0.11) for the scales of Miller et al. (2005). However, Colodro et al. (2018) analyzed the sex differences based on the normative sample of the official Spanish adaptation of the NEO PI-R and, given that this normative sample comprises people evaluated in personnel selection processes, the effect of social desirability common in this type of process may have cancelled out or attenuated the sex differences.

In fact, when comparing the means in the scales of Miller et al. (2005) obtained in the present study from a sample of adult volunteers and those obtained by Colodro et al. (2018) from the personnel selection sample of the official Spanish adaptation of the NEO PI-R, it seems clear that the norms of Colodro et al. (2018) for the PD scales show an important effect of social desirability. In the first place, and consistent with the previous scientific literature on the differences between volunteer samples and personnel selection samples (Sanz & García-Vera, 2009), in 60% of the scales of Miller et al. (2005), the mean scores of the Spanish adult volunteers in the present study exceeded of the norms of Colodro et al. (2018) by 0.50 or 0.80 standard deviation units, which is considered a medium or large effect size. In fact, in absolute value, the mean differences for all the scales ($d =$ 0.61) had a medium effect size. As might be expected, this indicated a more socially favorable

image for the norms based on a recruitment sample because, except for Histrionic and Obsessive-Compulsive PDs, the mean levels of the remaining eight PDs were markedly higher in the sample of Spanish volunteers. In addition, the scores on the scales of Miller et al. (2005) obtained by the Spanish adult volunteers in this study are similar to those obtained by volunteers from the general population of the United States, France and Belgium-Netherlands (on average, only 0.25, 0.23 and 0.11 standard deviations of difference, respectively). Secondly, using the same criterion as the cut-off score to identify individuals with possible PDs (a score equal to or greater than T 65 on the scales of Miller et al., 2005), when this cut-off score was based on the norms of Colodro et al. (2005), these researchers found very high prevalence rates of PDs in a non-clinical sample (e.g., 70% of the participants had a PD). However, when that cut-off score was based on the normative data from the present study, the prevalence rates of PDs in the sample of Spanish volunteers in this study were much lower (e.g., only 34.9% of the participants had a PD).

    Given these differences between the norms obtained by Colodro et al. (2018) based on the selection sample of personnel of the official Spanish adaptation of the NEO PI-R and the means obtained in the present study, it seems justified to develop specific norms for adult Spanish volunteers that allow the use of the PD scales of Miller et al. (2005) and of Costa and McCrae et al. (2005) in research contexts with volunteers or in situations where social desirability is presumed to be low, such as some clinical or educational contexts. In addition, given the sex differences, it also seemed justified to draw up specific differentiated norms for men and women. These norms are shown in Table 5. On this basis, an Excel spreadsheet was developed that allows the scores on both sets of PD scales to be automatically obtained from the direct or raw scores on the NEO PI-R facets (see supplementary material).

    A further and valuable issue is to establish the most appropriate cut-off score to identify the possible presence of a PD in the general population from these scales. Miller et al. (2005) recommended using a cut-off score greater than or equal to T 65, but Costa and McCrae (2005) recommended using BR scores. As indicated by the results of the present study, the former strategy could overestimate the presence of PDs because the scores on the scales follow a normal distribution and, therefore, by definition, approximately 6.68% of adults in the general population will obtain or exceed a T score of 65. This means that approximately 6.68% of adults would have each PD. This prevalence appears excessive according to the results of epidemiological studies (Winsper et al., 2020). The second strategy is limited by the validity of the prevalence rates of PDs in the population from which the BR scores are defined. Although this study used the results of the meta-analysis of epidemiological studies by Winsper et al. (2020) to estimate these prevalence rates, none of the studies in this meta-analysis had been conducted with samples from the general Spanish population. Future epidemiological studies conducted in Spain should fill this gap and provide adequate estimates of the prevalence rates of PDs in the general Spanish population with which to obtain more appropriate BR scores.

Another possibility would be to conduct a study with a sample of the Spanish population using a structured diagnostic interview as a criterion, and analyzing the receiver operating characteristic curves (ROC or predictive performance curves) to graphically represent the sensitivity and specificity for each of the direct scores of the scales and, consequently, select the direct score that offers the best diagnostic performance as the cut-off score for each scale (see Sanz, 2007). Meanwhile, the BR scores calculated in the present study could tentatively be the best estimate of the most appropriate cut-off scores, as they do not appear to overestimate the prevalence of PDs based on the limited epidemiological data currently available.

The results of this study regarding the sex differences in the PD scales of Miller et al. (2005) and Costa and McCrae et al. (2005), besides justifying obtaining separate norms for males and females, are consistent with the sex differences found in much of the scientific literature on the prevalence of PDs in the general population (Oltmanns & Powers, 2012). This literature suggests that paranoid, schizoid, schizotypal, antisocial, narcissistic, and obsessive-compulsive PDs are diagnosed more frequently in males and, conversely, borderline, histrionic, and dependent PDs are diagnosed more frequently in females (Oltmanns & Powers, 2012). Consistently, in this study and on both sets of scales, males scored significantly higher on paranoid, schizoid, antisocial, and narcissistic PDs, and females on borderline, histrionic, and dependent PDs. In addition, on one of the two sets of scales, males also scored significantly higher on schizotypal and obsessive-compulsive PDs. These sex differences in PDs seem to reflect the sex differences existing in normal personality dimensions and traits (Costa et al., 2001; Kaiser, 2019), and their explanation has to do with a multitude of factors that interact in a complex way and that are beyond the objectives of the present study, but that deserve in-depth investigation in future studies aimed specifically at clarifying them (Kaiser, 2019; Schmitt et al., 2017).

The results and conclusions of this study should be considered in light of its limitations. Among them, one limitation concerns the procedure for selecting the sample of participants. This incidental sample was recruited with the "snowball" technique and completed the NEO PI-R voluntarily and with guarantees of confidentiality as part of an investigation. Given the inherent limitations of this type of non-probability sampling, the degree to which the sample of participants thus obtained is representative of the Spanish adult population is questionable. However, in a variable as important as age, the profile of the sample of participants in this study involving three large age groups (18-29 years, 30-49 years, and 50 years and over) was very similar to that found in the Spanish population (Sanz & García-Vera, 2009). Nonetheless, it is obvious that the use of a random sample of participants belonging proportionately to different Spanish geographical regions would have greatly improved their representativeness and, therefore, the gen eralisation of their results. Another limitation of the study involves the convergent and discriminant validation strategy used, as both sets of PD scales, that of Miller et al. (2005) and that of Costa and McCrae (2005), are based on the facets of the NEO PI-R. Future

studies should include measures obtained with other instruments specifically designed and validated to measure PDs, such as the Exploratory Personality Questionnaire-III (*Cuestionario exploratorio de personalidad-III*, CEPER-III; Caballo et al., 2011), as convergent and discriminant validity criteria. Future studies should also examine which set of scales or which scale of each set is the most useful to identify people with PDs or to examine the factors, correlates or consequences of PDs.

Despite these limitations, the results of the present study provide adequate evidence of the internal consistency reliability and the convergent and discriminant validity in the general Spanish population of the measures of the PD scales based on the facets of the NEO PI-R proposed by Miller et al. (2005) and Costa and McCrae (2005). Likewise, the study provides norms and cut-off scores for these scales in the general Spanish population that complement those developed by Colodro et al. (2018) because although the latter are suitable for situations in which high social desirability is suspected (i.e., contexts of job selection or promotion or some forensic contexts), those of the present study would be appropriate for situations in which low or no social desirability is presumed (i.e., some clinical or educational contexts). These norms and cut-off scores would help identify people who probably have a PD according to the DSM-5, an identification that should be corroborated with a subsequent diagnostic interview, and they would also help to improve the clinical description of PDs considered as maladaptive variants of normal personality dimensions and traits.

References

American Psychiatric Association. (1995). *DSM-IV. Manual diagnóstico y estadístico de los trastornos mentales* [Diagnostic and statistical manual of mental disorders] (4th Ed.). Masson. (Orig. 1994).

American Psychiatric Association. (2014). *Manual diagnóstico y estadístico de los trastornos mentales. DSM-5* [Diagnostic and statistical manual of mental disorders. DSM-5] (5th Ed.). Editorial Médica Panamericana. (Orig. 2013).

Caballo, V. E., Guillén, J. L., Salazar, I. C., & Irurtia, M. J. (2011). Estilos y trastornos de personalidad: Características psicométricas del "Cuestionario exploratorio de personalidad-III" (CEPER-III) [Personality styles and disorders: Psychometric characteristics of the Personality Exploratory Questionnaire-III (CEPER-III)]. *Behavioral Psychology/Psicología Conductual, 19*(2), 277-302.

Colodro, J., López-García, J. J., Mezquita, L., Colodro-Conde, L., Ibáñez, M. I., Edo, S., Villa, E., & Ortet, G. (2018). Five-factor model of personality disorders: Spanish normative data and validation. *Anales de Psicología, 34*(2), 264-273. doi: 10.6018/analesps.34.2.289271

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory (NEO FFI). Professional manual*. Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1999). *Inventario de Personalidad NEO Revisado (NEO PI-R). Inventario NEO reducido de cinco factores (NEO FFI). Manual* [Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory (NEO FFI). Manual]. TEA Ediciones.

Costa, P. T., Jr., & McCrae, R. R. (2005). A five-factor model perspective on personality disorders. In S. Strack (Ed.), *Handbook of personology and psychopathology* (pp. 257-270). John Wiley & Sons.

Costa, P. T., Jr., & McCrae, R. R. (2008). *Inventario de personalidad NEO revisado (NEO PI-R). Inventario NEO reducido de cinco factores (NEO-FFI). Manual* [Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory (NEO FFI). Manual] (3rd rev. and expanded ed.). TEA Ediciones.

Costa, P. T., Jr., & McCrae, R. R. (2009). The five-factor model and the NEO Inventories. In J. N. Butcher (Ed.), *Oxford handbook of personality assessment* (pp. 299-322). Oxford University Press.

Costa, P. T., Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology, 81*(2), 322-331. doi: 10.1037/0022-3514.81.2.322

Hernández, A., Ponsoda, V., Muñiz, J., Prieto, G., & Elosua, P. (2016). Revisión del modelo para evaluar la calidad de los tests utilizados en España [Assessing the quality of tests in Spain: Revision of the Spanish test review model]. *Papeles del Psicólogo, 37*(3), 192-197.

John, O. P. (2021). History, measurement, and conceptual elaboration of the big-five trait taxonomy: The paradigm matures. In O. P. John y R. W. Robins (Eds.), *Handbook of personality: Theory and research* (pp. 35-82). Guilford.

Kaiser, T. (2019). Nature and evoked culture: sex differences in personality are uniquely correlated with ecological stress. *Personality and Individual Differences, 148*, 67-72. doi: 10.1016/j.paid.2019.05.011

Lynam, D. R., & Widiger, T. A. (2001). Using the five-factor model to represent the DSM-IV personality disorders: An expert consensus approach. *Journal of Abnormal Psychology, 110*(3), 401-412. doi: 10.1037/0021-843X.110.3.401

McCrae, R. R., & Costa, P. T., Jr. (2003). *Personality in adulthood. A five factor theory perspective* (2nd Ed.). Guilford.

Miller, J. D., Bagby, R. M., Pilkonis, P. A., Reynolds, S. K., & Lynam, D. R. (2005). A simplified technique for scoring DSM-IV personality disorders with the five-factor model. *Assessment, 12*(4), 404-415. doi: 10.1177/1073191105280987

Miller, J. D., & Lynam, D. R. (2008). Dependent personality disorder: Comparing an expert generated and empirically derived five-factor model personality disorder count. *Assessment, 15*(1), 4-15. doi: 10.1177/1073191107306095

Miller, J. D., Lynam, D. R., Rolland, J.-P., De Fruyt, F., Reynolds, S. K., Pham-Scottez, A., Baker, S. R., & Bagby, R. M. (2008). Scoring the DSM-IV personality disorders using the five-factor model: development and validation of normative scores for North American, French, and Dutch-Flemish samples. *Journal of Personality Disorders, 22*(5), 433-450. doi: 10.1521/pedi.2008.22.5.433

O'Connor, B. P. (2017). Robustness. In T. A. Widiger (Ed.), *The Oxford handbook of the five factor model* (pp. 151-171). Oxford University Press.

Oltmanns, T. F., & Powers, A. D. (2012). Gender and personality disorders. In T. A. Widiger (Ed.), *The Oxford handbook of personality disorders* (pp. 206-218). Oxford University Press.

Samuel, D. B., & Widiger, T. A. (2004). Clinicians' personality descriptions of prototypic personality disorders. *Journal of Personality Disorders, 18*(3), 286-308. doi: 10.1521/pedi.18.3.286.35446

Samuel, D. B., & Widiger, T. A. (2008). A meta-analytic review of the relationships between the five-factor model and DSM-IV-TR personality disorders: A facet level analysis. *Clinical Psychology Review, 28*(8), 1326-1342. doi: 10.1016/j.cpr.2008.07.002

Sanz, J. (2007). Algunos problemas con la utilización de la adaptación española del Inventario clínico multiaxial de Millon-II (MCMI-II) con fines diagnósticos [Some troubles with the use of the Spanish adaptation of Millon Clinical Multiaxial Inventory-II (MCMI-II) for diagnostic purposes]. *Clínica y Salud, 18*(3), 287-304.

Sanz, J. (2018). La evaluación de la personalidad mediante los inventarios NEO [Personality assessment with the NEO inventories]. In J. L. Graña Gómez, M. E. de la Peña Fernández, & J. M. Andreu Rodríguez (Eds.), *Manual docente de psicología clínica legal y forense* [Teaching manual on legal and forensic clinical psychology] (Vol. 1; 3rd Ed., pp. 1-74). Universidad Complutense de Madrid.

Sanz, J., & García-Vera, M. P. (2009). Nuevos baremos para la adaptación española del Inventario de Personalidad NEO Revisado (NEO PI-R): Fiabilidad y datos normativos en voluntarios de la población general [New norms for the Spanish adaptation of the Revised NEO Personality Inventory (NEO PI-R): Reliability and normative data in volunteers from the general population]. *Clínica y Salud, 20*(2), 131-144.

Sanz, J., Gil, F., García-Vera, M. P., & Barrasa, Á. (2008). Needs and cognition/behavior patterns at work and the big five: An assessment of the Personality and Preference Inventory-Normative (PAPI-N) from the perspective of the five-factor model. *International Journal of Selection and Assessment, 16*(1), 46-58. doi: 10.1111/j.1468-2389.2008.00408.x

Sanz, J., Silva, F., & Avia, M.D. (1999). La evaluación de la personalidad desde el modelo de los "cinco grandes": El "Inventario de cinco factores NEO" de Costa y McCrae [Personality assessment from the "Big Five" model: Costa and McCrae's "NEO Five Factor Inventory"]. In F. Silva (Ed.), *Avances en evaluación psicológica* (pp. 171 234). Promolibro.

Sanz-García, A., García-Vera, M. P., & Sanz, J. (2023). Is it time to replace the Big Five personality model? Factorial structure of the NEO PI-R in a community sample of Spanish adults. *The Journal of General Psychology*. Advance online publication. doi: 10.1080/00221309.2023.2261136

Schmitt, D. P., Long, A. E., McPhearson, A., O'Brien, K., Remmert, B., & Shah, S. H. (2017). Personality and gender differences in global perspective. *International Journal of Psychology, 52*(Suppl. 1), 45-56. doi: 10.1002/ijop.12265

Volkert, J., Gablonski, T., & Rabung, S. (2018). Prevalence of personality disorders in the general adult population in Western countries: Systematic review and meta-analysis. *The British Journal of Psychiatry, 213*(6), 709-715. doi: 10.1192/bjp.2018.202

Watson, D., & Clark, L. A. (2020). Personality traits as an organizing framework for personality pathology. *Personality and Mental Health, 14*(1), 51-75. doi: 10.1002/pmh.1458

Widiger, T. A., & Costa, P. T., Jr. (Eds.). (2013). *Personality disorders and the five-factor model of personality* (3rd Ed.). American Psychological Association.

Widiger, T. A., Gore, W. L., Crego, C., Rojas, S. L., & Oltmanns, J. R. (2017). Five factor model and personality disorder. In T. A. Widiger (Ed.), *The Oxford handbook of the five factor model* (pp. 449-478). Oxford University Press.

Widiger, T. A., Trull, T. J., Clarkin, J. F., Sanderson, C., & Costa, P. T., Jr. (2002). A description of the DSM-IV personality disorders with the five-factor model of personality. In P. T. Costa Jr., & T. A. Widiger (dirs.), *Personality disorders and the five-factor model of personality* (2nd Ed., pp. 89-99). American Psychological Association.

Winsper, C., Bilgin, A., Thompson, A., Marwaha, S., Chanen, A., Singh, S., Wang, A., & Furtado, V. (2020). The prevalence of personality disorders in the community: A global systematic review and meta-analysis. *The British Journal of Psychiatry, 216*(2), 69-78. doi: 10.1192/bjp.2019.166

Appendix 1

Norms for the personality disorder (PD) scales by Miller et al. (2005) derived from the NEO PI-R for males in the general population

PC	Paranoid	Schizoid	Schizotypal	Antisocial	Borderline	Histrionic	Narcissistic	Avoidant	Dependent*	Obses-com	PC
1	78	59	68	167	89	127	123	97	34	161	1
2	85	68	72	172	90	133	128	105	41	168	2
5	92	80	76	185	97	154	137	116	47	174	5
10	102	87	83	191	104	162	145	124	52	184	10
15	108	92	87	202	108	166	151	129	59	191	15
20	114	98	89	210	112	173	155	132	63	199	20
25	118	100	91	213	117	176	160	136	65	206	25
30	122	104	93	215	120	181	164	139	68	212	30
35	124	109	96	221	122	186	165	143	70	214	35
40	127	111	98	225	125	189	167	146	72	218	40
45	129	114	100	229	127	193	171	152	74	221	45
50	131	118	102	235	129	196	175	157	78	226	50
55	134	122	105	238	132	200	178	160	80	230	55
60	136	125	107	241	136	205	181	165	83	235	60
65	140	128	110	245	137	209	183	167	86	238	65
70	144	131	112	250	142	214	187	171	90	243	70
75	147	136	114	254	145	218	190	175	92	248	75
80	151	139	118	260	148	222	193	178	97	253	80
85	156	143	121	267	153	228	199	181	101	257	85
90	163	147	126	274	160	235	209	187	106	263	90
95	173	158	131	289	169	248	226	195	115	275	95
98	194	173	139	299	179	261	239	206	127	283	98
99	220	186	156	314	194	266	253	211	139	287	99

Notes: PC= percentile; OBSES-COM= Obsessive-Compulsive PD scale. *The Dependent PD Scale is the updated scale of Miller and Lynam (2008).

Appendix 2

Norms for the personality disorder (PD) scales by Miller et al. (2005) derived from the NEO PI-R for females in the general population

PC	Paranoid	Schizoid	Schizotypal	Antisocial	Borderline	Histrionic	Narcissistic	Avoidant	Dependent*	Obses-com	PC
1	78	59	68	167	89	127	123	97	34	161	1
2	85	68	72	172	90	133	128	105	41	168	2
5	92	80	76	185	97	154	137	116	47	174	5
10	102	87	83	191	104	162	145	124	52	184	10
15	108	92	87	202	108	166	151	129	59	191	15
20	114	98	89	210	112	173	155	132	63	199	20
25	118	100	91	213	117	176	160	136	65	206	25
30	122	104	93	215	120	181	164	139	68	212	30
35	124	109	96	221	122	186	165	143	70	214	35
40	127	111	98	225	125	189	167	146	72	218	40
45	129	114	100	229	127	193	171	152	74	221	45
50	131	118	102	235	129	196	175	157	78	226	50
55	134	122	105	238	132	200	178	160	80	230	55
60	136	125	107	241	136	205	181	165	83	235	60
65	140	128	110	245	137	209	183	167	86	238	65
70	144	131	112	250	142	214	187	171	90	243	70
75	147	136	114	254	145	218	190	175	92	248	75
80	151	139	118	260	148	222	193	178	97	253	80
85	156	143	121	267	153	228	199	181	101	257	85
90	163	147	126	274	160	235	209	187	106	263	90
95	173	158	131	289	169	248	226	195	115	275	95
98	194	173	139	299	179	261	239	206	127	283	98
99	220	186	156	314	194	266	253	211	139	287	99

Notes: PC= percentile; OBSES-COM= Obsessive-Compulsive PD scale. *The Dependent PD Scale is the updated scale of Miller and Lynam (2008).

Appendix 3

Norms for the personality disorder (PD) scales by Costa and McCrae (2005) derived from the NEO PI-R for males in the general population

PC	Paranoid	Schizoid	Schizotypal	Antisocial	Borderline	Histrionic	Narcissistic	Avoidant	Dependent	Obses-com	PC
1	26	19	94	64	61	98	69	57	99	85	1
2	30	25	100	68	65	101	73	63	107	87	2
5	34	30	107	73	72	123	78	70	120	96	5
10	38	35	112	84	79	128	83	78	132	104	10
15	41	37	116	91	87	132	85	82	137	107	15
20	44	40	119	96	90	138	88	86	140	110	20
25	45	42	121	99	92	142	89	89	142	114	25
30	46	45	123	101	96	145	91	91	145	116	30
35	48	47	125	103	101	148	94	94	148	118	35
40	49	48	127	106	103	151	95	98	149	121	40
45	51	51	130	110	105	153	97	101	152	123	45
50	53	53	133	112	108	155	99	104	155	124	50
55	55	55	135	114	110	157	100	108	157	126	55
60	56	56	137	117	112	161	102	110	158	128	60
65	57	58	139	120	115	163	104	112	160	130	65
70	60	61	143	123	118	165	106	115	162	134	70
75	62	63	146	127	121	168	108	119	166	137	75
80	64	65	149	133	125	171	109	123	169	139	80
85	68	68	155	137	130	175	112	127	173	144	85
90	73	71	159	144	135	184	115	133	178	149	90
95	80	79	166	149	146	193	122	141	187	156	95
98	93	86	177	166	158	204	133	146	200	168	98
99	98	93	185	170	175	207	138	162	200	172	99

Note: PC= percentile; OBSES-COM= Obsessive-Compulsive PD scale.

Appendix 4

Norms for the personality disorder (PD) scales by Costa and McCrae (2005) derived from the NEO PI-R for females in the general population

PC	Paranoid	Schizoid	Schizotypal	Antisocial	Borderline	Histrionic	Narcissistic	Avoidant	Dependent	Obses-com	PC
1	26	19	94	64	61	98	69	57	99	85	1
2	30	25	100	68	65	101	73	63	107	87	2
5	34	30	107	73	72	123	78	70	120	96	5
10	38	35	112	84	79	128	83	78	132	104	10
15	41	37	116	91	87	132	85	82	137	107	15
20	44	40	119	96	90	138	88	86	140	110	20
25	45	42	121	99	92	142	89	89	142	114	25
30	46	45	123	101	96	145	91	91	145	116	30
35	48	47	125	103	101	148	94	94	148	118	35
40	49	48	127	106	103	151	95	98	149	121	40
45	51	51	130	110	105	153	97	101	152	123	45
50	53	53	133	112	108	155	99	104	155	124	50
55	55	55	135	114	110	157	100	108	157	126	55
60	56	56	137	117	112	161	102	110	158	128	60
65	57	58	139	120	115	163	104	112	160	130	65
70	60	61	143	123	118	165	106	115	162	134	70
75	62	63	146	127	121	168	108	119	166	137	75
80	64	65	149	133	125	171	109	123	169	139	80
85	68	68	155	137	130	175	112	127	173	144	85
90	73	71	159	144	135	184	115	133	178	149	90
95	80	79	166	149	146	193	122	141	187	156	95
98	93	86	177	166	158	204	133	146	200	168	98
99	98	93	185	170	175	207	138	162	200	172	99

Note: PC= percentile; OBSES-COM= Obsessive-Compulsive PD scale.