TPM®

STRUCTURAL ANALYSIS AND MEASUREMENT INVARIANCE REGARDING THE SEX OF THREE BRIEF VERSIONS OF THE BIG FIVE INVENTORY IN MEXICAN GENERAL POPULATION

SERGIO DOMINGUEZ-LARA
UNIVERSIDAD PRIVADA NORBERT WIENER, LIMA, PERÙ

YOLANDA CAMPOS-USCANGA
UNIVERSIDAD VERACRUZANA, XALAPA, VERACRUZ, MÉXICO

Assessing personality with short scales is becoming relevant in different fields of psychology. Therefore, the purpose of this study is to analyze the internal structure and reliability of three short versions of the Big Five Inventory in a sample of 500 Mexican citizens (60.8% men; M_{age} = 23.58). The German (BFI-15; 15 items), Peruvian (BFI-15p; 15 items), and Brazilian (BFI-20; 20 items) versions were analyzed using exploratory structural equation modeling. The Peruvian version (BFI-15p) showed a simpler factorial structure and factor loadings with acceptable magnitudes, as well as adequate reliability coefficients. Likewise, the BFI-15p was invariant between men and women, with women scoring greater Agreeableness, Neuroticism, and Openness than men. The implications of these findings and recommendations for future studies are discussed.

Keywords: Personality; Big Five factors; Validity; Reliability; Measurement invariance; Sex differences.

Correspondence concerning this article should be addressed to Yolanda Campos-Uscanga, Instituto de Salud Pública, Universidad Veracruzana, Luis Castelazo Ayala s/n, Col. Industrial Ánimas, C.P. 91190, Xalapa, Veracruz, Mexico. Email: ycampos@uv.mx or *yolycampos@gmail.com*

Personality is assumed as a set of qualities that defines the way a person feels, thinks, and behaves (Schacter et al., 2011). Although individual differences are extremely diverse, most of them do not have major impacts on daily life. Common response patterns, encompassing specific constant traits, have been identified, giving rise to the theory of personality traits (Goldberg, 1990). Personality traits are consistent and enduring ways of reacting to the environment, and these responses tend to be similar to different stimuli. They are in constant interaction with the environment and even determine a person's search for stimuli.

Since its origin, the 5-factor personality model has shown to be reliable in multiple studies and useful to explain behavior, as personality relates to other constructs. According to this model, there are five big factors of personality: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness (BFF; McCrae & Costa, 2008).

The first factor, Extraversion, involves people with high enthusiasm and energy from external means, both material and social. In addition, they enjoy interacting with people, are enthusiastic, and carry out many activities in such a way that they are always focused on doing something to avoid boredom or routine (John et al., 2008; McCrae & Costa, 1999). On the other hand, the Agreeableness trait implies a prosocial and community orientation toward others. There is a concern for social harmony, so their social relations are usually harmonious. They are perceived as kind, generous, and trustworthy people. Instead, Conscientiousness is related to impulse control based on the social standards of one's in-group. Highly

conscious people tend to focus on accomplishing tasks and goals, such as thinking before acting, delaying gratification, following standards and rules, and planning, organizing, and prioritizing tasks. As for Neuroticism, it is characterized by experiencing negative emotions such as anxiety, sadness, and tension. Likewise, people exhibit low tolerance to stress and aversive stimuli, in addition to perceiving ordinary situations as threatening. This negative emotion derived from some situations may last for long periods of time, which is why bad moods are frequent. Finally, Openness implies the willingness to experiment, the search for adventure, imagination, and curiosity. There is an enjoyment in new things, with an interest in art and aesthetics, and people tend to be creative, but also risk takers (John et al., 2008; McCrae & Costa, 1999).

This BFF model has been shown to be exhaustive and applicable in different cultures (McCrae & John, 1992). There are several instruments for its measurement, among which the NEO Personality Inventory-Revised (NEO-PI-R; Costa & McCrae, 1999) and the Big Five Inventory (BFI; John et al., 1991) stand out, and the latter having among its main advantages its effectiveness, the application in less time, and that its items are shorter and easier to understand (John & Srivastava, 1999).

The biological differences between men and women, especially in terms of reproduction, are well accepted; however, there are controversies in psychological aspects that require reliable measurement tools to be able to make these comparisons. Among the initial controversies are the underlying causes of these differences, but regardless of whether these causes are developmental or sociocultural, understanding different personalities is essential to understand the changes throughout the lifespan (Schmitt et al., 2017).

Differences According to Sex in Personality Factors

Overall, the differences in personality traits between men and women can be explained by biological, cultural, and psychological adaptation processes (Schmitt et al., 2017). For example, women achieve higher scores than men in Extraversion (Mac Giolla & Kajonius, 2019; Schmitt et al., 2008), while there is evidence that they have similar levels in Agreeableness (Bunnett, 2020; Chapman et al., 2007; De Bolle et al., 2015; Dominguez-Lara et al., 2019; Marsh et al., 2013; Schmitt et al., 2008; Soto et al., 2011; Weisberg et al., 2011) and Neuroticism (Bunnett, 2020; Chapman et al., 2007; De Bolle et al., 2015; Dominguez-Lara et al., 2019; Mac Giolla & Kajonius, 2019; Marsh et al., 2013; Schmitt et al., 2008; Weisberg et al., 2011). Regarding Conscientiousness, women also score higher than men (Dominguez-Lara et al., 2019; Mac Giolla & Kajonius, 2019; Schmitt et al., 2008), although in some cases there are no differences (Chapman et al., 2007; Soto et al., 2011; Weisberg et al., 2011). Finally, equivalence is mostly found in Openness between groups (Chapman et al., 2007; Soto et al., 2011), but in one study it was found that women score higher (Mac Giolla & Kajonius, 2019).

Big Five Inventory and Its Use in General Population

The BFI is one of the most widely used scales for studying personality traits since it has some comparative advantages regarding its commercial counterparts, such as the NEO-PI-R, especially in terms of its extension and its costs for the average user; the BFI has only 44 items and is free to distribute. Although it has been used in many specific groups, it is also used in the general population due to its association with behaviors and health outcomes, and significant predictive power. For example, during the COVID-19 pandemic, Extraversion, Agreeableness, Conscientiousness, and Openness were found to be inversely associated with generalized anxiety and depressive symptoms, while Neuroticism was directly associated (Nikčević et

al., 2021). Likewise, greater Conscientiousness and less Neuroticism are associated with better health and more healthy behaviors (Rochefort et al., 2019).

The BFI has also been used to study eating habits and physical activity in the general population. Orthorexia nervosa has been found to be negatively associated with Agreeableness (Strahler et al., 2020), whereas Openness has been reported as related to physical activity (Pristyna et al., 2022). Similarly, in Japanese women it was found that the extraverted personality trait is associated to greater physical activity (Satoh et al., 2022). These behaviors are associated with obesity, which has also shown to be positively related to Extraversion and Conscientiousness and negatively related to Neuroticism, the latter acting as a protective factor. On the other hand, under the hypothesis that personality and metabolism should be correlated if they function as an integrated unit, a study was carried out in young adults where it was found that less extraverted individuals had a higher resting metabolic rate, suggesting a compensation of energy between personality and basal metabolism (Bergeron et al., 2021).

The continuous interaction that occurs today among countries and cultures facilitates opportunities, but also challenges related to acceptance and openness to diversity. The BFI has also been used for these researches in the general population, finding that higher scores in the traits of Openness and Agreeableness, as well as lower scores in Neuroticism and Conscientiousness, are associated to greater openness to diversity (Han & Pistole, 2017).

Big Five Inventory: Short Versions and Previous Evidence

The BFI's extension does not intrinsically represent a difficulty when the objective is to assess personality, but when it is part of a study that also considers other constructs, the used scales' total extension could cause fatigue in the examinee and, consequently, subtract validity from the answers. In this context, two short versions of 10 items (BFI-10; Rammstedt & John, 2007) and 15 items (BFI-15; Gerlitz & Schupp, 2005) emerged, which assess each dimension with two and three items, respectively, but it is necessary to highlight some aspects.

As for the BFI-10 (Rammstedt & John, 2007), as it is made up of a direct item and an inverse one in relation to the dimension it assesses, it is susceptible to the effect of the method (Lance et al., 2010), and even of socially desirable responses (Gomes & Gjikuria, 2017). In the same way, its association with the extended version was not corrected due to the existence of items in common, which can artificially increase the magnitude of the association (Levy, 1967), and since there is no reliability report, it is not possible to quantify the measurement error tolerated by that score, so its use could be questionable.

On the other hand, the BFI-15's construction (Gerlitz & Schupp, 2005) was based on the combination of principal component analysis (estimation method) and varimax rotation, which are not recommended in psychometric studies oriented toward self-reports (Lloret-Segura et al., 2014; Watkins, 2018). Furthermore, like the BFI-10, the association with the long version was not corrected (Levy, 1967), although its reliability estimates using the α coefficient were acceptable (between .50 and .73; see Ponterotto & Charter, 2009).

Regarding the reliability analysis by internal consistency, this is usually based on the calculation and report of the α coefficient, whose magnitude depends, among other things, on the number of items, so it is usual that in this type of short version (two or three items per dimension) the coefficients are relatively low, so it is necessary to expand this information with an independent indicator of the number of items: average interitem correlation (r_{ij}; Clark & Watson, 1995).

Psychometric evidence from contexts similar to the original indicates adequate psychometric properties (Courtois et al., 2020; Guido et al., 2015; Hahn et al., 2012; Rammstedt, 2007), while in samples with characteristics different from the previous ones favorable indicators on the internal structure or reliability of the BFI-10 and BFI-15 are not obtained (Balgiu, 2018; Dominguez-Lara & Merino-Soto, 2018a; Kim et al., 2010; Kunnel et al., 2019; Lovik et al., 2017; Pejić et al., 2014; Steyn & Ndofirepi, 2022), and although short versions are known to obtain smaller α magnitudes than full scales (e.g., Lang et al., 2011; Rammstedt & Beierlein, 2014), sometimes the magnitudes are too low for both the α (< .56; Chapman & Elliot, 2019; Lovik et al., 2017) and the r_{ij} (< .32; Chapman & Elliot, 2019).

For this reason, and considering some already mentioned limitations, alternative versions of 10 and 15 items (BFI-10p and BFI-15p; Dominguez-Lara & Merino-Soto, 2018a) were created from the BFI's Spanish version (Benet-Martínez & John, 1998). Its elaboration was based on a semiconfirmatory approach based on the procrustean factor rotation (Browne, 1972). To structure them, items with direct wording were selected, and with greater affinity to the construct, considering both the magnitude of the factor loadings as well as the congruence coefficient (Lorenzo-Seva & ten Berge, 2006) and the factorial simplicity index (Fleming & Merino, 2005), also reporting the empirical equivalence with the extended version corrected for the presence of items in common (Levy, 1967). Subsequently, favorable psychometric evidence was found in university students from Peru, Chile, and Mexico regarding the validity of its internal structure, its association with other variables, and its reliability indicators (Dominguez-Lara & Merino-Soto, 2018b; Dominguez-Lara et al., 2022, 2023), although the analysis of measurement invariance across sex has not yet been explored.

Recently, Gouveia et al. (2021) proposed a 20-item version (BFI-20) based on the BFI's Brazilian version. For this purpose, they used the extraction method by principal axes, eigenvalue greater than unity to determine the number of dimensions, and varimax rotation, while to verify congruence with the original structure they used procrustes rotation. It should be noted that the equivalence with the extended version was not verified either, and loads above .30 were considered significant. Finally, the reported reliability was low in some cases, such as in Conscientiousness (α = .56; ω = .55), and Openness (α = .60; ω = .61), being more acceptable in Agreeableness (α = .69; ω = .64), Extraversion (α = .72; ω = .73), and Neuroticism (α = .69; ω = .72).

The Present Study

The aim of this research is to analyze the psychometric properties of three short versions of BFI: Peruvian (Dominguez-Lara & Merino-Soto, 2018a), German (Gerlitz & Schupp, 2005), and Brazilian (Gouveia et al., 2021) in the general Mexican population. The use of brief instruments is increasingly widespread in the field of research as well as in the applied field, and although there is psychometric evidence that could support its use in certain contexts, the focus of attention has been limited to samples made up only of university students, without including postgraduate students, workers, unemployed, and people dedicated to housework who are located in the same age group. This group has particular characteristics and, sometimes, the psychometric properties of the assessment instruments are different from those found in university students since these are not representative of the general population (Preti et al., 2013; Whisman & Judd, 2016) and the manifestations of the personality dimensions could be different depending on the experiences lived by the people (Mõttus & Rozgonjuk, 2021), and even show changes before vital events (Cobb-Clark & Schurer, 2012). On the other hand, it would be important to have an instrument that briefly and accurately assesses BFFs due to their importance in the workplace (Zell & Lesick, 2022) or clinical setting (Dash et al., 2019), just to mention the most prominent within psychology.

Material and Methods

Participants and Procedure

The sample was nonprobabilistic, and the sampling was intentional. As inclusion criteria, we considered being between 18 and 40 years of age, and not having a previously diagnosed personality disorder.

A total of 500 Mexican citizens (60.8% men) between 18 and 40 years old participated (M_{age} = 23.58, SD_{age} = 5.35); 66.40% were university and postgraduate students, 26.80% were salaried or independent workers, 2.40% were dedicated to housework, 2.20% were unemployed, and 2.20% did not respond. Regarding the demographic data, most of them stated that they did not have a partner (85.4%) or children (85.8%) when answering questions. Furthermore, no differences were found between men and women in relation to having a partner (ϕ = .10) or children (ϕ = .08), although small differences were found in occupation (ϕ = .11), since there were a higher number of male university students compared to female university students, and a small age difference (d = .31) in favor of women.

Data collection was carried out in July 2022 in public and private spaces for recreation and physical exercise in the city of Xalapa (Veracruz, Mexico). Adults who came to these spaces between 7:00 and 18:00 were approached to invite them to participate in the study, and those who agreed to participate signed the informed consent at that time and answered the questionnaire in printed format. The responses have no missing data in the sociodemographic questionnaire except for occupation, or in self-reports.

Measures

Initially participants responded to a sociodemographic data questionnaire in which they reported on their sex, age, relationship (yes or not), number of children, and current occupation. The Big Five Inventory-15p (Dominguez-Lara & Merino-Soto, 2018a) is a self-report measure that assesses BFFs with 15 items (three items per dimension). The items come from the Spanish version of BFI (Benet-Martínez & John, 1998), which are scaled in Likert format (from 1 = *strongly disagree* to 5 = *strongly agree*).

The Big Five Inventory-15 (Gerlitz & Schupp, 2005), similar to the BFI-15p, assesses BFFs with three items per dimension scaled in Likert format (from 1 = *strongly disagree* to 5 = *strongly agree*). For the present study, this version was constructed from the items of the BFI in its Spanish version (Benet-Martínez & John, 1998).

The Big Five Inventory-20 (Gouveia et al., 2021) assesses the BFFs with four items per dimension, which are rated from *strongly disagree* (1) to *strongly agree* (5). For the present study, this version was constructed from the items of the BFI in its Spanish version (Benet-Martínez & John, 1998).

Ethical Considerations

This study is part of a larger project entitled "Emotional intelligence, personality traits, emotional state, training spaces and type of physical exercise in adults," approved by Comité de Ética en Investigación del Instituto de Salud Pública de la Universidad Veracruzana (Universidad Veracruzana's Institute of Public Health's Research Ethics Committee) (Registration CEI-ISP-R11/2022). The participants read an informed consent prior to the application of the scale and the requirements established in the Declaration of Helsinki were met.

Analysis Procedure

Structural Analysis

The multivariate outliers were computed with the Mahalanobis Distances and a *p*-value less than .001 indicates the presence of an outlier (Kline, 2016). The structural analysis was based on exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009; Marsh et al., 2014) with the WLSMV estimation method and geomin rotation (ε = .50; Marsh et al., 2009, 2011), based on polychoric matrices, as conducted in previous studies (Dominguez-Lara & Merino-Soto, 2018b; Dominguez-Lara et al., 2022).

The models were evaluated based on their fit indices, as well as on the magnitude of the factor loadings. In this sense, the comparative fit index (CFI) and Tucker Lewis index (TLI) (> .90; Marsh et al., 2014), and the root-mean-square error of approximation (RMSEA) were assessed, being considered appropriate if confidence interval's (CI) lower limit is less than .05 and the upper limit is less than .10 (West et al., 2012), and the weighted root-mean-square residual (WRMR) as well (≤ 1.00; DiStefano et al., 2018). On the other hand, factor loadings greater than .50 were expected taking into account the number of items per dimension (Dominguez-Lara, 2018a), and those greater than .30 were considered significant secondary loads. Additionally, in order to quantify the relevance of secondary loadings, the factor simplicity index (FSI; Fleming & Merino, 2005) was reported, where magnitudes greater than .70 (Lara et al., 2021) would indicate that the item receives predominant influence of a single factor.

The construct's reliability was estimated with the ω coefficient (> .70), and the scores' reliability was estimated with the α coefficient (> .65; Ponterotto & Charter, 2009) and the average interitem correlation (r_{ij} > .20; Clark & Watson, 1995). The ω and α coefficients were compared between men and women considering the difference's CI interval (CI_{diff}; Moreta-Herrera et al., 2021), that is, if the CI includes zero, there are no differences in terms of the reliability estimate between groups; on the other hand, the r_{ij} was compared between groups using the *q* coefficient (Cohen, 1992): less than .10, insignificant difference; between .10 and .30, small; between .30 and .50, moderate; and greater than .50, large.

Measurement Invariance and Group Comparison

Subsequently, evidence of measurement invariance between men and women was obtained through a multigroup factorial analysis with the best model among the three available, considering configural invariance (equivalence of the internal structure), weak (equivalence of the factor loadings), strong (equivalence of thresholds), and strict (equivalence of residuals) (Pendergast et al., 2017). The degree of measurement invariance was generally assessed, considering the variation of their fit indices (CFI and RMSEA), that is, there is evidence of invariance if ΔCFI > –.01 and ΔRMSEA ≤ .02 (Chen, 2007). Regarding the comparison between groups, the difference in the BFFs was assessed according to the *d* coefficient (Cohen, 1992) under the effect size approach: less than .20, insignificant difference; between .20 and .50, small; between .50 and .80, moderate; and greater than .80, large.

Software

To execute the ESEM and the measurement invariance, the Mplus Version 7 program (Muthén & Muthén, 1998/2015) was used. To calculate Cohen's *d* and *q*, a specific module was used (Dominguez-Lara, 2018b).

RESULTS

The analysis of the multivariate outliers showed that the group of men had a high proportion of data that are not considered multivariate outliers (BFI-15 = 99.01%; BFI-20 = 98.36%; BFI-15= 98.03%), and in the women's group, it was also found that a high proportion were not outliers (BFI-15 = 97.45%; BFI-20 = 98.47%; BFI-15 = 97.45%). The proportion of multivariate outliers was not considered problematic (Magalhães et al., 2014), so we worked with the entire sample, since these data represent possible observations within the general population, so it would help to better represent that group (Baião et al., 2015).

Regarding the internal structure, the BFI-15 obtained favorable fit indices in the group of men — CFI = .97; TLI = .91; RMSEA [90% CI] = .06 [.04, .07]; WRMR = .45 —, although nearly half of the items had unacceptable factor loadings (seven of 15), and 12 of the 15 items were not factorially simple (FSI < .70; see Table 1). The analysis could not be carried out in the group of women due to model identification problems associated with Item 25 (Conscientiousness).

TABLE 1
ESEM analysis and factor simplicity of BFI-15 (Gerlitz & Schupp, 2005) in Mexican men

	F1	F2	F3	F4	F5	FSI
F1: Extraversion						
Item 1	.44	–.10	.16	–.07	.35	.52[a]
Item 6	–.58	.08	.13	.05	.36	.65[a]
Item 43	.73	.02	.03	–.07	.14	.95
F2: Agreeableness						
Item 22	.25	–.45	–.31	.27	.25	.28[a]
Item 28	.20	.24	–.29	–.01	.10	.41[a]
Item 37	.05	.65	–.06	.07	.24	.83
F3: Conscientiousness						
Item 3	–.01	–.32	.66	.09	.40	.56[a]
Item 25	–.05	–.11	–.61	.16	.36	.65[a]
Item 29	.10	.42	.34	.04	.29	.42[a]
F4: Neuroticism						
Item 9	–.10	.27	–.10	–.62	.49	.51[a]
Item 26	–.09	.24	.09	.60	.16	.73
Item 38	–.23	–.07	–.17	.49	.15	.63[a]
F5: Openness						
Item 5	.24	.11	.43	.00	.32	.48[a]
Item 17	–.01	.31	.16	.09	.24	.48[a]
Item 20	.28	.29	.15	.14	.29	.27[a]
F1	1					
F2	.15	1				
F3	.21	.27	1			
F4	–.02	–.05	–.15	1		
F5	.15	.28	.14	.22	1	

Note. ESEM = exploratory structural equation modeling; BFI = Big Five Inventory. The numbering of the items corresponds to the original version by Benet-Martínez and John (1998). FSI = factor simplicity index; [a] = FSI < .70.

On the BFI-20 (Gouveia et al., 2021), the fit indices were adequate in the men group — CFI = .98; TLI = .95; RMSEA [90% CI] = .05 [.04, .06]; WRMR = .52 —, with acceptable factor loadings in 14 of the 20 items, and 15 of the 20 items were factorially simple (FSI > .70; see Table 2). In the group of women, the fit was also acceptable — CFI = .94; TLI = .88; RMSEA [90% CI] = .07 [.06, .09]; WRMR = .58 —, with adequate factor loadings in 15 of the 20 items, and with 16 factorially simple items (FSI > .70; see Table 3). In addition, both groups share the most representative items in three of the five dimensions: Agreeableness, Item 37, "É amável, tem consideração pelos outros [Is considerate and kind to almost everyone]"; Neuroticism, Item 15, "Fica tenso com frequência [Can be tense]"; and Openness, Item 23, "É inventivo, criativo [Is inventive]".

TABLE 2
ESEM analysis and factor simplicity of BFI-20 (Gouveia et al., 2021) in Mexican men

	F1	F2	F3	F4	F5	FSI
F1: Extraversion						
Item 1	.41	–.11	.08	.04	.17	.73
Item 11	.42	.24	.01	–.20	.22	.47[a]
Item 32	.46	.15	.01	–.02	.10	.84
Item 43	.99	–.03	-.03	.08	–.13	.97
F2: Agreeableness						
Item 7	.03	.45	.23	–.02	.20	.63[a]
Item 28	.08	.34	–.29	.04	.16	.49[a]
Item 37	.03	.68	.07	.04	.06	.97
Item 41	.17	.63	.15	.00	.02	.86
F3: Conscientiousness						
Item 3	–.01	–.32	.56	.05	.25	.60[a]
Item 14	–.06	.17	.76	–.01	–.04	.93
Item 21	.04	.04	.64	.01	.03	.99
Item 29	.16	.25	.53	.05	.05	.71
F4: Neuroticism						
Item 15	.04	–.05	–.01	.82	.00	.99
Item 26	–.06	.18	.16	.57	.10	.78
Item 30	.14	–.13	.07	.58	–.04	.87
Item 38	–.11	.06	–.12	.64	–.01	.91
F5: Openness						
Item 5	.18	–.04	.25	–.08	.46	.62[a]
Item 20	.06	.09	–.13	.13	.75	.91
Item 23	.12	–.08	.02	–.10	.77	.94
Item 36	–.16	.22	.13	.09	.60	.74
F1	1					
F2	.29	1				
F3	.32	.39	1			
F4	–.13	.03	.02	1		
F5	.56	.47	.62	.03	1	

Note. ESEM = exploratory structural equation modeling; BFI = Big Five Inventory. The numbering of the items corresponds to the original version by Benet-Martínez and John (1998). FSI = factor simplicity index; [a] = FSI < .70.

TABLE 3
ESEM analysis and factor simplicity of BFI-20 (Gouveia et al., 2021) in Mexican women

	F1	F2	F3	F4	F5	FSI
F1: Extraversion						
Item 1	.77	.08	.10	.10	–.18	.89
Item 11	.21	.08	.22	–.26	.29	.26[a]
Item 32	.26	.20	.05	–.14	.39	.47[a]
Item 43	.55	.10	–.04	–.04	.24	.77
F2: Agreeableness						
Item 7	.04	.61	.05	–.01	.10	.95
Item 28	.08	.23	–.07	–.01	.04	.77
Item 37	.02	.83	.01	.09	–.10	.97
Item 41	.18	.57	.08	–.12	–.03	.83
F3: Conscientiousness						
Item 3	.21	–.22	.44	.10	–.11	.54[a]
Item 14	.07	.01	.62	.09	.04	.95
Item 21	–.10	–.10	.75	–.05	–.01	.95
Item 29	–.02	.30	.59	–.02	.02	.76
F4: Neuroticism						
Item 15	.14	–.07	.05	.89	–.02	.96
Item 26	–.00	–.01	.09	.72	.14	.94
Item 30	.15	.05	–.15	.61	.05	.85
Item 38	–.36	.07	.06	.64	.07	.71
F5: Openness						
Item 5	.18	–.12	.06	.03	.73	.89
Item 20	–.15	.13	.02	.19	.78	.86
Item 23	.12	–.08	.01	–.07	.82	.96
Item 36	.15	.18	.07	.09	.37	.61[a]
F1	1					
F2	.23	1				
F3	.29	.36	1			
F4	–.12	–.10	.02	1		
F5	.31	.42	.39	–.06	1	

Note. ESEM = exploratory structural equation modeling; BFI = Big Five Inventory. The numbering of the items corresponds to the original version by Benet-Martínez and John (1998). FSI = factor simplicity index; [a] = FSI < .70.

Regarding the BFI-15p, the model in the group of men showed favorable fit indices — CFI = .99; TLI = .97; RMSEA [90% CI] = .04 [.02, .06]; WRMR = .37 —, with acceptable factor loadings, except in Items 11 and 23, and evidencing that 13 of the 15 items are factorially simple (FSI > .70; see Table 4).

In the women's group, the fit was also acceptable — CFI = .97; TLI = .93; RMSEA [90% CI] = .06 [.03, .08]; WRMR = .40 —, with adequate factor loadings in 11 of the 15 items, and 14 items factorially simple (FSI > .70; see Table 5). Additionally, both groups share the most representative items in Extraversion, Item 43: "Es extrovertido, sociable [Is outgoing, sociable]"; Agreeableness, Item 37: "Es considerado

y amable con casi todo el mundo [Is considerate and kind to almost everyone]"; Conscientiousness, Item 21: "Persevera hasta terminar el trabajo [Perseveres until the task is finished]"; and Neuroticism, Item 15: "Con frecuencia se pone tenso [Can be tense]"; while in Openness the most representative item differs between men, Item 17: "Valora lo artístico, lo estético [Values artistic, aesthetic experiences]," and women, Item 39: "Es educado en arte, música o literatura [Is sophisticated in art, music, or literature]."

As for reliability of both the construct (ω coefficient) and scores (α coefficient), acceptable magnitudes were found in most cases (see Table 6). Differences between men and women in BFI-20 were found in the α coefficient in two of the five dimensions (Agreeableness and Conscientiousness), while differences were only found in favor of men in the ω coefficient and the r_{ij} in Extraversion and Conscientiousness, respectively. Regarding the BFI-15p, a sole difference was found in Openness in favor of women in the ω coefficient and r_{ij} in comparison to men.

TABLE 4
ESEM analysis and factor simplicity of BFI-15p (Dominguez-Lara & Merino-Soto, 2018a) in Mexican men

	F1	F2	F3	F4	F5	FSI
F1: Extraversion						
Item 1	.52	.12	–.05	.06	.01	.91
Item 11	.48	.23	.12	–.17	.09	.62[a]
Item 43	.79	–.02	–.04	.05	.08	.98
F2: Agreeableness						
Item 7	.07	.62	.17	–.01	.00	.90
Item 37	–.07	.78	–.09	.01	.06	.97
Item 41	.13	.74	.00	.02	–.01	.96
F3: Conscientiousness						
Item 14	–.13	.14	.69	.00	.10	.89
Item 21	.07	–.05	.71	.01	.03	.98
Item 34	.04	.05	.52	.02	–.03	.98
F4: Neuroticism						
Item 4	–.16	–.01	–.02	.61	.08	.90
Item 15	.05	.11	–.01	.77	–.10	.95
Item 30	.14	–.08	.06	.63	.06	.90
F5: Openness						
Item 17	–.23	.06	.10	.08	.79	.88
Item 23	.34	–.01	.30	–.05	.33	.41[a]
Item 39	.14	–.04	–.22	–.06	.75	.86
F1	1					
F2	.37	1				
F3	.29	.57	1			
F4	–.17	–.12	–.04	1		
F5	.32	.47	.36	–.04	1	

Note. ESEM = exploratory structural equation modeling; BFI = Big Five Inventory. The numbering of the items corresponds to the original version by Benet-Martínez and John (1998). FSI = factor simplicity index; [a] = FSI < .70.

TABLE 5
ESEM analysis and factor simplicity of BFI-15p (Dominguez-Lara & Merino-Soto, 2018a) in Mexican women

	F1	F2	F3	F4	F5	FSI
F1: Extraversion						
Item 1	.41	.08	.00	.16	.09	.77
Item 11	.34	.09	.13	–.28	.18	.40[a]
Item 43	.96	–.04	–.03	.07	–.03	.99
F2: Agreeableness						
Item 7	.10	.70	.11	–.01	–.10	.93
Item 37	–.12	.82	–.10	.13	.11	.91
Item 41	.06	.61	.06	–.15	–.05	.90
F3: Conscientiousness						
Item 14	.02	.11	.47	.03	.14	.84
Item 21	–.03	–.08	.89	–.02	–.04	.99
Item 34	.07	.15	.36	.08	–.01	.75
F4: Neuroticism						
Item 4	.01	.05	–.14	.63	–.03	.94
Item 15	–.04	–.06	.12	.88	.05	.97
Item 30	.09	.02	.00	.60	–.01	.97
F5: Openness						
Item 17	–.02	–.04	.10	.24	.70	.85
Item 23	.24	.00	–.01	–.13	.55	.77
Item 39	–.06	.03	–.05	–.10	.89	.98
F1	1					
F2	.44	1				
F3	.33	.30	1			
F4	–.08	–.10	–.04	1		
F5	.40	.42	.40	–.10	1	

Note. ESEM = exploratory structural equation modeling; BFI = Big Five Inventory. The numbering of the items corresponds to the original version by Benet-Martínez and John (1998). FSI = factor simplicity index; [a] = FSI < .70.

TABLE 6
Reliability differences regarding sex

	r_{ij}			α			ω		
	M	W	*q*	M	W	CI_{diff}	M	W	CI_{diff}
BFI-20									
Extraversion	.35	.34	.01	.68	.67	–.08, .11	.68	.52	.05, .30
Agreeableness	.31	.25	.05	.64	.57	.05, .19	.61	.66	–.16, .05
Conscientiousness	.38	.26	.11	.71	.58	.03, .25	.72	.69	–.06, .12
Neuroticism	.37	.43	.05	.70	.75	–.13, .03	.75	.81	–.13, .00
Openness	.47	.45	.02	.78	.76	–.05, .09	.75	.78	–.10, .04
BFI-15p									
Extraversion	.37	.32	.04	.64	.59	–.07, .18	.63	.62	–.10, .14
Agreeableness	.47	.41	.05	.73	.68	–.04, .16	.76	.76	–.07, .08
Conscientiousness	.37	.28	.09	.64	.53	–.02, .25	.68	.62	–.05, .18
Neuroticism	.39	.43	.03	.66	.69	–.13, .07	.72	.76	–.12, .04
Openness	.35	.45	.11	.62	.71	–.19, .02	.67	.76	–.18, –.00

Note. BFI = Big Five Inventory. r_{ij} = average interitem correlation; M = men; W = women; *q* = Cohen's *q*; CI_{diff} = difference's confidence interval.

According to the factorial parameters (factor loadings and FSI) and reliability coefficients (and their differences), the best BFF model was the BFI-15p, and the measurement invariance was performed on this version. In this sense, the fit indices' variation supports the configural, weak, strong, and strict invariance (see Table 7), which makes it possible to compare the BFFs between men and women and, in turn, to carry out the analysis of the factor structure of the BFI-15p with the total sample.

TABLE 7
Invariance analysis regarding sex

	CFI	RMSEA	90% CI	WRMR	ΔCFI	ΔRMSEA
Configural	.98	.05	[.03, .06]	.54		
Weak	.99	.03	[.01, .05]	.74	.01	–.02
Strong	.98	.04	[.02, .05]	.91	–.01	.00
Strict	.97	.04	[.03, .05]	1.04	–.01	.01

Note. CFI = comparative fit index; RMSEA = root-mean-square error of approximation; CI = confidence interval; WRMR = weighted root-mean-square residual.

The model in the total sample obtained favorable fit indices, CFI = .99; TLI = .97; RMSEA [90% CI] = .04 [.02, .05]; WRMR = .38, with acceptable factor loadings, except in the Items 11, 34, and 17, and evidencing that 13 of the 15 items are factorially simple (FSI > .70; see Table 8).

TABLE 8
ESEM analysis and factor simplicity of BFI-15p (Dominguez-Lara & Merino-Soto, 2018a) in the total Mexican sample

	F1	F2	F3	F4	F5	FSI
F1: Extraversion						
Item 1	.50	.11	–.05	.13	.04	.86
Item 11	.45	.14	.11	–.24	.10	.61[a]
Item 43	.83	–.03	–.03	.09	.03	.98
F2: Agreeableness						
Item 7	.08	.69	.12	.00	–.06	.94
Item 37	–.08	.81	–.12	.05	.08	.95
Item 41	.08	.68	.05	–.04	–.03	.97
F3: Conscientiousness						
Item 14	–.08	.13	.63	.02	.10	.90
Item 21	.01	–.09	.78	.00	.02	.98
Item 34	.06	.09	.44	.04	–.03	.92
F4: Neuroticism						
Item 4	–.11	.04	–.08	.62	.04	.93
Item 15	.02	.00	.08	.82	–.03	.99
Item 30	.14	–.03	.02	.62	.01	.94
F5: Openness						
Item 17	.28	.02	.17	–.08	.41	.55[a]

(table 8 continues)

Table 8 (continued)

	F1	F2	F3	F4	F5	FSI
(*F5: Openness*)						
Item 23	–.17	.01	.10	.14	.79	.90
Item 39	.08	–.04	–.15	–.08	.78	.93
F1	1					
F2	.41	1				
F3	.34	.52	1			
F4	–.17	–.09	–.02	1		
F5	.38	.48	.41	–-.00	1	

Note. ESEM = exploratory structural equation modeling; BFI = Big Five Inventory. The numbering of the items corresponds to the original version by Benet-Martínez and John (1998). FSI = factor simplicity index; [a] = FSI < .70.

Finally, as shown in Table 9, women scored greater Agreeableness, Neuroticism, and Openness than men (d > .20), while no differences in Extraversion and Conscientiousness were found.

TABLE 9
Comparison between men and women in Big Five factors

Dimension	Group	*M*	*SD*	*d*
Extraversion	Men	10.82	2.76	0.12
	Women	10.50	2.83	
Agreeableness	Men	12.01	2.41	0.28
	Women	12.66	2.12	
Conscientiousness	Men	11.52	2.30	0.17
	Women	11.91	2.34	
Neuroticism	Men	7.74	2.76	0.28
	Women	8.57	3.14	
Openness	Men	10.76	2.58	0.21
	Women	11.33	2.81	

DISCUSSION

The short version's psychometric study of any measurement instrument represents a challenge since, in addition to the natural decrease of some indicators (e.g., reliability coefficients), the construct's breadth and empirical representativeness must be reflected in favorable parameters (e.g., factor proper loadings). In this sense, and since the short versions of the BFI are widely used instruments, it was deemed relevant to explore their psychometric characteristics.

Similar to preliminary studies (e.g., Dominguez-Lara & Merino-Soto, 2018a), the internal structure of the short German version (Gerlitz & Schupp, 2005) was not supported in the studied sample; despite having acceptable fit indices, many of the items showed unacceptable factor complexity and relatively low factor loadings. This finding reinforces the idea that it is not advisable to consider the fit indices as the only

source of information to provide evidence of validity in relation to the internal structure of an instrument. In addition, as demonstrated in other studies (Steyn & Ndofirepi, 2022), it is necessary to analyze the instruments' internal structure, especially if the cultural context where it was generated differs from the context in which it will be administered. Similarly, even though the Brazilian version (Gouveia et al., 2021), showed some strengths, its psychometric performance was below the BFI-15p, taking into account the number of factorially simple items and the magnitude of its factor loadings.

On the other hand, the BFI-15p (Dominguez-Lara & Merino-Soto, 2018a) showed more solid indicators in terms of its configuration as in the group of men 13 of the 15 items were presented as factorially simple, while in the group of women the number was 14 of the 15 items. It can be seen that the construct's empirical representativeness was adequate considering the magnitude of the factor loadings, and the predominance of factorially simple items reflects the specificity of each one of them to assess the dimension for which it was built, minimizing the influence of secondary factors. On the other hand, the most representative items (markers) of each dimension are consistent with the definition of each factor, since Extraversion speaks of the individual's ability to socialize, Item 43: "Es extrovertido, sociable [Is outgoing, sociable]," while Agreeableness is based on the warmth of interpersonal relationships, Item 37: "Es considerado y amable con casi todo el mundo [Is considerate and kind to almost everyone]." On the other hand, the Conscientiousness marker is associated with tenacity and perseverance, Item 21: "Persevera hasta terminar el trabajo [Perseveres until the task is finished]," while the Neuroticism marker focuses on the experience of negative emotions, Item 15: "Con frecuencia se pone tenso [Can be tense]." Finally, in the Openness dimension there are discrepancies as the marker differs between men, Item 17: "Valora lo artístico, lo estético [Values artistic, aesthetic experiences]", and women, Item 39: "Es educado en arte, música o literatura [Is sophisticated in art, music, or literature]," although at a conceptual level the differences are minor since both statements are linked to art, and from an empirical point of view the difference between factor loadings was insignificant between the groups (< .10). As for reliability, the indicators are acceptable in all cases, and the magnitudes do not differ between men and women. These results are consistent with and reinforce the findings of other studies carried out on university students in Peru (Dominguez-Lara & Merino-Soto, 2018b), Mexico (Dominguez-Lara et al., 2022), and Chile (Dominguez-Lara et al., 2023), where the 5-dimensional structure presented adequate factor loadings, as well as acceptable factorial simplicity.

Lastly, the BFI-15p is shown to be invariant between men and women, that is, this version can be used to make fair comparisons between groups. Sex differences in the BFFs are interesting as it is the first study of this kind carried out in a Mexican population. There are some coincidences with the results found in other populations, but important differences are also found, which confirms the relevance of having an adapted instrument to deepen research in this regard.

Previous studies show that women achieve higher scores than men in Extraversion (e.g., Mac Giolla & Kajonius, 2019); however, in the present work similar scores were observed between men and women in Extraversion. Apparently, the expression of this trait occurs in a similar way in men and women, which turns out to be relevant as it is linked to the expression of healthy behaviors (Nikčević et al., 2021). On the other hand, the trait of Conscientiousness is controversial since while some studies found higher scores in women (e.g., Mac Giolla & Kajonius, 2019), others did not report sex differences (e.g., Soto et al., 2011). The latter are consistent with the results of the present study and this trait has also been linked to healthy behaviors and mental health (Nikčević et al., 2021; Rochefort et al., 2019). In the present study, higher scores were found for women in three personality traits that, according to previous studies, did not show differences between men and women: Neuroticism (e.g., Bunnett, 2020), Agreeableness (e.g., Dominguez-Lara et al., 2019), and Openness (e.g., Soto et al., 2011). Also, a study reported higher scores of Openness in women than men in 22 countries (Mac Giolla & Kajonius, 2019), coinciding with these results.

Therefore, Neuroticism is especially relevant given that it is a trait that has been linked to negative effects on health, which could explain the higher prevalence of many physical and mental illnesses in women (Nikčević et al., 2021; Rochefort et al., 2019). On the other hand, the highest scores in Agreeableness and Openness in women coincide with the cultural characteristics and traditional sex roles still present in Mexico. Previous studies indicate that there is a strong correlation between personality differences by sex in a country and its gender equality index (Mac Giolla & Kajonius, 2019), which is one of the great challenges facing the country.

This study has some strengths, such as the use of ESEM, which is increasingly used in the field of psychometrics, as well as the exploration of measurement invariance according to sex, which was not carried out in preliminary studies for methodological reasons. In the same way, it allows the reader to expose the need to complement the report and results' interpretation considering other aspects relevant to the assessed model in addition to the fit indices (Dominguez-Lara, 2016).

Before mentioning the findings' practical implications, it should be noted that the BFI-15p version shows robustness in terms of its internal structure, since, although it was previously applied to Mexican university students obtaining favorable results (Dominguez-Lara et al., 2022), it continues to present good psychometric performance compared to other versions (German and Brazilian). For this reason, having a short scale to measure personality in the general Mexican population will strengthen basic and applied research as it can be used in massive surveys (e.g., Rammstedt, 2007). In addition, given the complex reality of multifactorial diseases, it is necessary to identify personality characteristics that are related to health damage or protective behaviors. Therefore, the use of BFI-15p in future studies that seek to relate personality to relevant events for physical and mental health is very promising, and since it is invariant according to sex, its use would be free of bias.

As for the limitations, since the study is focused on the internal structure, for reasons of space it was not possible to present other validity evidences. Likewise, although the reliability coefficients are acceptable considering the number of items per dimension (Ponterotto & Charter, 2009) and it is possible to use this version in the field of basic research, its implementation is not recommended when dealing with situations that involve decision making regarding the examinee (e.g., job interviews) due to the amount of measurement error they tolerate. From the chosen sample's point of view, it is clear that this is not representative of the general Mexican population since only one region was considered, although this would not have a substantial impact on the results because Xalapa is a city that brings together people from many parts of the country so, although they reside in the same space, there is a lot of diversity, in addition to the fact that the cultural differences between regions are not very marked. However, having the participation of people who attend public spaces, it is probable that there is a self-selection bias since they would be individuals with certain personality characteristics (Feng et al., 2022). Finally, while the validation targeted adults aged 18 to 40, it is important to assess its functioning in older adults. Additionally, since the majority of participants were university students, future studies should focus on examining the scale's properties in a population with a lower prevalence of students.

Hence, it is concluded that the BFI-15p has a solid internal structure and acceptable reliability coefficients in Mexican men and women from the general population. For future studies we recommend replicating the study in other Mexican regions, in order to consolidate the version, and using stratified sampling by age and occupation. On the other hand, the similarity with previous studies conducted in the Spanish-speaking population, and the discrepancies with the Brazilian version, suggest some type of invariability in the BFI-15p in countries with similar cultural and language characteristics. Therefore, it would be advisable to explore the cross-cultural invariance in future studies. Lastly, given that the most representative items of

each dimension remain stable between groups, it would be convenient to explore the psychometric benefits of a 5-item version of the BFI.

Acknowledgements

The authors thank Campus para la Cultura, las Artes y el Deporte (Campus for Culture, Arts, and Sports) of Universidad Veracruzana and the Estadio Xalapeño "Heriberto Jara Corona" from the Government of the State of Veracruz, for access facilities to carry out this research.

References

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 397-438. https://doi.org/10.1080/10705510903008204

Baião, R., Gilbert, P., McEwan, K., & Carvalho, S. (2015). Forms of Self-Criticising/Attacking & Self-Reassuring Scale: Psychometric properties and normative study. *Psychology and Psychotherapy*, *88*(4), 438-452. https://doi.org/10.1111/papt.12049

Balgiu, B. A. (2018). The psychometric properties of the Big Five Inventory-10 (BFI-10) including correlations with subjective and psychological well-being. *Global Journal of Psychology Research: New Trends and Issues*, *8*(2), 61-69. https://doi.org/10.18844/gjpr.v8i2.3434

Benet-Martínez, V., & John, O. P. (1998). *Los Cinco Grandes* across cultures and ethnic groups: Multitrait-multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology*, *75*(3), 729-750. https://doi.org/10.1037/0022-3514.75.3.729

Bergeron, P., Pagé, A., & Trempe, M. (2021). Integrating humans into pace-of-life studies: The Big Five personality traits and metabolic rate in young adults. *PloS One*, *16*(4), Article e0248876. https://doi.org/10.1371/journal.pone.0248876

Browne, M. W. (1972). Oblique rotation to a partially specified target. *British Journal of Mathematical and Statistical Psychology*, *25*(2), 207-212. https://doi.org/10.1111/j.2044-8317.1972.tb00492.x

Bunnett, E. R. (2020). Gender differences in perceived traits of men and women. In B. J. Carducci, C. S. Nave, A. Di Fabio, D. H. Saklofske, & C. Stough (Eds.), *The Wiley encyclopedia of personality and individual differences: Personality processes and individual differences* (pp. 179-184). John Wiley & Sons. https://doi.org/10.1002/9781119547174.ch207

Chapman, B. P., Duberstein, P. R., Sörensen, S., & Lyness, J. M. (2007). Gender differences in Five Factor Model personality traits in an elderly cohort. *Personality and Individual Differences*, *43*(6), 1594-1603. https://doi.org/10.1016/j.paid.2007.04.028

Chapman, B. P., & Elliot, A. J. (2019). Brief report: How short is too short? An ultra-brief measure of the Big Five personality domains implicates "agreeableness" as a risk for all-cause mortality. *Journal of Health Psychology*, *24*(11), 1568-1573. https://doi.org/10.1177/1359105317720819

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464-504. https://doi.org/10.1080/10705510701301834

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309-319. https://doi.org/10.1037/1040-3590.7.3.309

Cobb-Clark, D. A., & Schurer, S. (2012). The stability of Big Five personality traits. *Economics Letters*, *115*(1), 11-15. https://doi.org/10.1016/j.econlet.2011.11.015

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159. https://doi.org/10.1037/0033-2909.112.1.155

Costa, P. T., Jr., & McCrae, R. R. (1999). *Manual NEO PIR, Inventario de Personalidad NEO Revisado y NEO-FFI, Inventario NEO reducido de Cinco Factores* [NEO PIR manual, revised NEO Personality Inventory and NEO-FFI, reduced NEO Five-Factor Inventory]. TEA Ediciones.

Courtois, R., Petot, J. M., Plaisant, O., Allibe, B., Lignier, B., Réveillère, C., Lecocq, G., & John, O. (2020). Validation française du Big Five Inventory à 10 items (BFI-10) [French validation of the 10-item Big Five Inventory]. *L'Encéphale*, *46*(6), 455-462. https://doi.org/10.1016/j.encep.2020.02.006

Dash, G. F., Slutske, W. S., Martin, N. G., Statham, D. J., Agrawal, A., & Lynskey, M. T. (2019). Big Five personality traits and alcohol, nicotine, cannabis, and gambling disorder comorbidity. *Psychology of Addictive Behaviors*, *33*(4), 420-429. https://doi.org/10.1037/adb0000468

De Bolle, M., De Fruyt, F., McCrae, R. R., Löckenhoff, C. E., Costa, P. T., Jr., Aguilar-Vafaie, M. E., Ahn, C.-k., Ahn, H.-n., Alcalay, L., Allik, J., Avdeyeva, T. V., Bratko, D., Brunner-Sciarra, M., Cain, T. R., Chan, W., Chittcharat, N., Crawford, J. T., Fehr, R., Ficková, E., ... Terracciano, A. (2015). The emergence of sex differences in personality traits in early adolescence: A cross-sectional, cross-cultural study. *Journal of Personality and Social Psychology*, *108*(1), 171-185. https://doi.org/10.1037/a0038497

DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the weighted root mean square residual: Evidence for trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 453-466. https://doi.org/10.1080/10705511.2017.1390394

Dominguez-Lara, S. (2016). Evaluación de modelos estructurales, más allá de los índices de ajuste [Evaluation of structural models, beyond fit indices]. *Enfermería Intensiva*, *27*(2), 84-85. https://doi.org/10.1016/j.enfi.2016.03.003

Dominguez-Lara, S. (2018a). Propuesta de puntos de corte para cargas factoriales: Una perspectiva de fiabilidad de constructo [Proposal for cut-offs for factor loadings: A construct reliability perspective]. *Enfermería Clínica*, *28*(6), 401-402. https://doi.org/10.1016/j.enfcli.2018.06.002

Dominguez-Lara, S. (2018b). Magnitud del efecto: Una guía rápida [Effect size: A quick guide]. *Educación Médica*, *19*(4), 251-254. https://doi.org/10.1016/j.edumed.2017.07.002

Dominguez-Lara, S., Campos-Uscanga, Y., & Valente, S. (2022). Análisis psicométrico de versiones breves del Big Five Inventory en universitarios mexicanos [Psychometric analysis of short versions of the Big Five Inventory in Mexican college students]. *Avaliação Psicológica*, *21*(2), 140-149. https://doi.org/10.15689/ap.2022.2102.20163.02

Dominguez-Lara, S., Carmona-Halty, M., Schulmeyer, M. K., Valente, S. N., & Lourenço, A. A. (2023). *Análisis psicométrico de tres versiones del Big Five Inventory-15 en una muestra de estudiantes universitarios chilenos: Estructura interna y asociación con bienestar subjetivo* [Psychometric analysis of three versions of the Big Five Inventory -15 in Chilean college students: internal structure and association with subjective well-being] [Manuscript submitted for publication].

Dominguez-Lara, S., & Merino-Soto, C. (2018a). Dos versiones breves del Big Five Inventory en universitarios peruanos: BFI-15p y BFI-10p [Two brief versions of the Big Five Inventory in Peruvian university students: BFI-15p and BFI-10p]. *Liberabit*, *24*(1), 81-96. https://doi.org/10.24265/liberabit.2018.v24n1.06

Dominguez-Lara, S., & Merino-Soto, C. (2018b). Estructura interna del BFI-10P y BFI-15P: Un estudio complementario con enfoque CFA y ESEM [Internal structure of BFI-10P and BFI-15P: A complementary study with CFA and ESEM approach]. *Revista Argentina de Ciencias del Comportamiento*, *10*(3), 22-34. https://doi.org/10.32348/1852.4206.v10.n3.21037

Dominguez-Lara, S., Prada-Chapoñan, R., & Moreta-Herrera, R. (2019). Diferencias de género en la influencia de la personalidad sobre la procrastinación académica en estudiantes universitarios peruanos [Gender differences in the influence of personality on academic procrastination in Peruvian college students]. *Acta Colombiana de Psicología*, *22*(2), 125-136. https://doi.org/10.14718/acp.2019.22.2.7

Feng, X., Astell-Burt, T., Standl, M., Flexeder, C., Heinrich, J., & Markevych, I. (2022). Green space quality and adolescent mental health: Do personality traits matter? *Environmental Research*, *206*, Article 112591. https://doi.org/10.1016/j.envres.2021.112591

Fleming, J. S., & Merino, C. (2005). Medidas de simplicidad y ajuste factorial: Un enfoque para la construcción y revisión de escalas derivadas factorialmente [Measures of simplicity and factorial fit: An approach for the evaluation of factorially derived scales]. *Revista de Psicología*, *23*(2), 252-266. https://doi.org/10.18800/psico.200502.002

Gerlitz J. Y., & Schupp, J. (2005). *Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP. Dokumentation der Instrumentenentwicklung BFI-S auf Basis des SOEP-Pretests 2005* [On the survey of Big Five-based personality traits in the SOEP. Documentation of the BFI-S instrument development based on the SOEP pretest 2005]. DIW Research, Notes 4.

Goldberg, L. R. (1990). An alternative "description of personality": The Big Five factor structure. *Journal of Personality and Social Psychology*, *59*(6), 1216-1229. https://doi.org/10.1037/0022-3514.59.6.1216

Gomes, C. M. A., & Gjikuria, J. (2017). Comparing the ESEM and CFA approaches to analyze the Big Five factors. *Avaliação Psicológica*, *16*(3), 261-267. https://doi.org/10.15689/ap.2017.1603.12118

Gouveia, V., Rodrigues, R., Oliveira, I. C., Gonçalves, M., Milfont, T., Coelho, G., Santos, W., de Medeiros, E. D., Soares, A. K., Monteiro, R., Andrade, J., Cavalcanti, T., Nascimento, B., & Gouveia, R. (2021). A short version of the Big Five Inventory (BFI-20): Evidence on construct validity. *Interamerican Journal of Psychology*, *55*(1), Article e1312. https://doi.org/10.30849/ripijp.v55i1.1312

Guido, G., Peluso, A. M., Capestro, M., & Miglietta, M. (2015). An Italian version of the 10-item Big Five Inventory: An application to hedonic and utilitarian shopping values. *Personality and Individual Differences*, *76*, 135-140. https://doi.org/10.1016/j.paid.2014.11.053

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory — Versions 4a and 54*. Institute of Personality and Social Research, University of California, Berkeley.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114-158). Guilford Press.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 102-138). Guilford Press.

Hahn, E., Gottschling, J., & Spinath, F. M. (2012). Short measurements of personality — Validity and reliability of the GSOEP Big Five Inventory (BFI-S). *Journal of Research in Personality*, *46*(3), 355-359. https://doi.org/10.1016/j.jrp.2012.03.008

Han, S., & Pistole, M. C. (2017). Big Five personality factors and facets as predictors of openness to diversity. *The Journal of Psychology*, *151*(8), 752-766. https://doi.org/10.1080/00223980.2017.1393377

Kim, S. Y., Kim, J. M., Yoo, J. A., Bae, K. Y., Kim, S. W., Yang, S. J., Shin, I. S., & Yoon, J. S. (2010). Standardization and validation of Big Five Inventory — Korean version (BFI-K) in elders. *Korean Journal of Biological Psychiatry*, *17*(1), 15-25.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Kunnel, J. R., Xavier, B., Waldmeier, A., Meyer, A., & Gaab, J. (2019). Psychometric evaluation of the BFI-10 and the NEO-FFI-3 in Indian adolescents. *Frontiers in Psychology*, *10*, Article 1057. https://doi.org/10.3389/fpsyg.2019.01057

Lance, C., Dawson, B., Birkelbach, D., & Hoffman, B. (2010). Method effects, measurement error, and substantive conclusions. *Organizational Research Methods*, *13*(3), 435-455. https://doi.org/10.1177/1094428109352528

Lang, F. R., John, D., Lüdtke, O., Schupp, J., & Wagner, G. G. (2011). Short assessment of the Big Five: Robust across survey methods except telephone interviewing. *Behavior Research Methods*, *43*(2), 548-567. https://doi.org/10.3758/s13428-011-0066-z

Lara, L., Monje, M. F., Fuster-Villaseca, J., & Dominguez-Lara, S. (2021). Adaptación y validación del Big Five Inventory para estudiantes universitarios chilenos [Adaptation and validation of the Big Five Inventory in Chilean university students]. *Revista Mexicana de Psicología*, *38*(2), 83-94.

Levy, P. (1967). The correction for spurious correlation in the evaluation of short-form tests. *Journal of Clinical Psychology*, *23*(1), 84-86. https://doi.org/10.1002/1097-4679(196701)23:1<84::aid-jclp2270230123>3.0.co;2-2

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada [Exploratory item factor analysis: A practical guide revised and updated]. *Anales de Psicología*, *30*(3), 1151-1169. https://doi.org/10.6018/analesps.30.3.199361

Lovik, A., Verbeke, G., & Molenberghs, G. (2017). Evaluation of a very short test to measure the Big Five personality factors on a Flemish sample. *Journal of Psychological and Educational Research*, *25*(2), 7-17.

Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *2*(2), 57-64. https://doi.org/10.1027/1614-2241.2.2.57

Mac Giolla, E., & Kajonius, P. J. (2019). Sex differences in personality are larger in gender equal countries: Replicating and extending a surprising finding. *International Journal of Psychology*, *54*(6), 705-711. https://doi.org/10.1002/ijop.12529

Magalhães, E., Salgueira, A., Gonzalez, A-J., Costa, J. J, Costa, M. J., Costa, P., & Pedroso de Lima, M. (2014). NEO-FFI: Psychometric properties of a Short Personality Inventory in Portuguese context. *Psicologia: Reflexão e Crítica*, *27*(4), 642-657. https://doi.org/10.1590/1678-7153.201427405

Marsh, H. W., Liem, G. A. D., Martin, A. J., Morin, A. J. S., & Nagengast, B. (2011). Methodological measurement fruitfulness of exploratory structural equation modeling (ESEM): New approaches to key substantive issues in motivation and engagement. *Journal of Psychoeducational Assessment*, *29*(4), 322-346. https://doi.org/10.1177/0734282911406657

Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, *10*, 85-110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 439-476. https://doi.org/10.1080/10705510903008220

Marsh, H. W., Nagengast, B., & Morin, A. J. S. (2013). Measurement invariance of Big-Five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology*, *49*(6), 1194-1218. https://doi.org/10.1037/a0026913

McCrae, R. R., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 139-153). Guilford Press.

McCrae, R. R., & Costa, P. T., Jr. (2008). Empirical and theoretical status of the five-factor model of personality traits. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment, Vol. 1. Personality theories and models* (pp. 273-294). Sage Publications, Inc. https://doi.org/10.4135/9781849200462.n13

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*(2), 175-215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

Moreta-Herrera, R., Dominguez-Lara, S., Sánchez-Guevara, S., López-Castro, J., & Molina-Narváez, M. J. (2021). Análisis multigrupo por sexo y fiabilidad del Cuestionario de Regulación Emocional (ERQ) en jóvenes ecuatorianos [Multigroup analysis by sex and reliability of the Emotional Regulation Questionnaire (ERQ) in Ecuadorians]. *Avaliação Psicológica*, *20*(2), 220-228. https://doi.org/10.15689/ap.2021.2002.19889.10

Mõttus, R., & Rozgonjuk, D. (2021). Development is in the details: Age differences in the Big Five domains, facets, and nuances. *Journal of Personality and Social Psychology*, *120*(4), 1035-1048. https://doi.org/10.1037/pspp0000276

Muthén, L. K., & Muthén, B. O. (2015). *Mplus user's guide*. Muthén & Muthén. (Original work published 1998).

Nikčević, A. V., Marino, C., Kolubinski, D. C., Leach, D., & Spada, M. M. (2021). Modelling the contribution of the Big Five personality traits, health anxiety, and COVID-19 psychological distress to generalised anxiety and depressive symptoms during the COVID-19 pandemic. *Journal of Affective Disorders*, *279*, 578-584. https://doi.org/10.1016/j.jad.2020.10.053

Pejić, M., Tenjović, L., & Knežević, G. (2014). Validacija upitnika ličnosti BFI-10-kratke forme inventara velikih pet [Validation of the BFI-10 questionnaire — Short version of the Bg Five Inventory]. *Primenjena Psihologija*, *7*(1), 45-92. https://doi.org/10.19090/pp.2014.1.45-92

Pendergast, L. L., von der Embse, N., Kilgus, S. P., & Eklund, K. R. (2017). Measurement equivalence: A non-technical primer on categorical multi-group confirmatory factor analysis in school psychology. *Journal of School Psychology*, *60*, 65-82. https://doi.org/10.1016/j.jsp.2016.11.002

Ponterotto, J. G., & Charter R. A. (2009). Statistical extensions of Ponterotto and Ruckdeschel's (2007) reliability matrix for estimating the adequacy of internal consistency coefficients. *Perceptual and Motor Skills*, *108*(3), 878-886. https://doi.org/10.2466/PMS.108.3.878-886

Preti, A., Vellante, M., Gabbrielli, M., Lai, V., Muratore, T., Pintus, E., Pintus, M., Sanna, S., Scanu, R., Tronci, D., Corrias, I., Petretto, D. R., & Carta, M. G. (2013). Confirmatory factor analysis and measurement invariance by gender, age and levels of psychological distress of the short TEMPS-A. *Journal of Affective Disorders*, *151*(3), 995-1002. https://doi.org/10.1016/j.jad.2013.08.025

Pristyna, G., Mahmudiono, T., Rifqi, M. A., & Indriani, D. (2022). The relationship between Big Five personality traits, eating habits, physical activity, and obesity in Indonesia based on analysis of the 5th wave Indonesia Family Life Survey (2014). *Frontiers in Psychology*, *13*, Article 881436. https://doi.org/10.3389/fpsyg.2022.881436

Rammstedt, B. (2007). The 10-item Big Five Inventory (BFI-10): Norm values and investigation of sociodemographic effects based on a German population representative sample. *European Journal of Psychological Assessment*, *23*(3), 193-201. https://doi.org/10.1027/1015-5759.23.3.193

Rammstedt, B., & Beierlein, C. (2014). Can't we make it any shorter? *Journal of Individual Differences*, *35*(4), 212-220. https://doi.org/10.1027/1614-0001/a000141

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203-212. https://doi.org/10.1016/j.jrp.2006.02.001

Rochefort, C., Hoerger, M., Turiano, N. A., & Duberstein, P. (2019). Big Five personality and health in adults with and without cancer. *Journal of Health Psychology*, *24*(11), 1494-1504. https://doi.org/10.1177/1359105317753714

Satoh, M., Sato, N., & Fujimura, A. (2022). Relationship between personality traits and health behavior among working women in Japan. *Preventive Medicine Reports*, *25*, Article 101691. https://doi.org/10.1016/j.pmedr.2021.101691

Schacter, D. L., Gilbert, D. T., & Wegner, D. M. (2011). *Psychology* (2nd ed.). Worth.

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*(1), 45-56. https://doi.org/10.1002/ijop.12265

Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology*, *94*(1), 168-182. https://doi.org/10.1037/0022-3514.94.1.168

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, *100*(2), 330-348. https://doi.org/10.1037/a0021717

Steyn, R., & Ndofirepi, T. M. (2022). Structural validity and measurement invariance of the short version of the Big Five Inventory (BFI-10) in selected countries, *Cogent Psychology*, *9*(1), Article 2095035. https://doi.org/10.1080/23311908.2022.2095035

Strahler, J., Haddad, C., Salameh, P., Sacre, H., Obeid, S., & Hallit, S. (2020). Cross-cultural differences in orthorexic eating behaviors: Associations with personality traits. *Nutrition*, *77*, Article 110811. https://doi.org/10.1016/j.nut.2020.110811

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, *44*(3), 219-246. https://doi.org/10.1177/0095798418771807

Weisberg, Y. J., DeYoung, C. G., & Hirsh, J. B. (2011). Gender differences in personality across the ten aspects of the Big Five. *Frontiers in Psychology*, *2*, Article 178. https://doi.org/10.3389/fpsyg.2011.00178

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209-234). Guilford Press.

Whisman, M. A., & Judd, C. M. (2016). A cross-national analysis of measurement invariance of the Satisfaction with Life Scale. *Psychological Assessment*, *28*(2), 239-244. https://doi.org/10.1037/pas0000181

Zell, E., & Lesick, T. L. (2022). Big Five personality traits and performance: A quantitative synthesis of 50+ meta-analyses. *Journal of Personality*, *90*(4), 559-573. https://doi.org/10.1111/jopy.12683