



Cultural Challenges of the Portuguese Adaptation of an IPIP English-Based Personality Inventory

Andrea Ritter^{1,2}, Hugo Esteves², and Aristides Ferreira¹

¹Department of Human Resources and Organizational Behaviour, ISCTE-IUL, Instituto Universitário de Lisboa, Lisbon, Portugal

²Department of People Science, EIKKO – HR Technology Solutions, Lda., Instituto Pedro Nunes, Coimbra, Portugal

Abstract: An instrument to measure job-related traits was developed and adapted to a Portuguese sample. This inventory, an adaptation of items from the International Personality Item Pool (Goldberg et al., 2006), measures a set of work-related facets over two studies ($N = 437$). Given the challenges in translating and adapting words and idioms from English to Portuguese, the focus was to ensure the scale remained gender-neutral and that its psychometric quality was enhanced. The findings validated 16 work-related personality scales under a three-factor structure, suggesting a new theoretical model for work-related personality traits in the Portuguese language and professional context. This research underscores the importance of exploring personality in the workplace, particularly within the Portuguese context with its distinct cultural nuances.

Keywords: personality, Big Five, facet models, work personality, gender-neutral

Personality has consistently been shown to be an important predictor of work-related behavior, especially within the model that denotes the Big Five personality traits. This model has served as a taxonomy (Goldberg, 1993) allowing research to grow, but since most of the studies are focused on the broader domains (Barrick & Mount, 1991), its predictive utility has been criticized (Morgeson et al., 2007). To increase predictive power in personality tests, two main methods have been endorsed: the use of narrow personality facet models (Dudley et al., 2006) and a contextualized measure using the frame-of-reference (FOR; Lievens et al., 2008). While many facet models under the Big Five domain have been emerging over the last 30 years, agreement on a perfect model is far from being reached (Costa & McCrae, 1998), especially if we account for the 274 scales available in the International Personality Item Pool (IPIP; Goldberg et al., 2006). In line with Schulze et al. (2021), for a facet model to be an enduring predictor of behaviors at work, it needs to be developed considering elements such as “time, action, target, and context” (Schulze et al., 2021, p. 369). This inventory was developed considering (1) a specific context through the use of frame-of-reference (FOR) “at work” (Hunthausen et al., 2003), (2) an active Portuguese working population, and (3) the current period

of time, bearing in mind the shifts in labor and skills demand (Deming, 2017). Accordingly, the purpose of this study is to adapt an IPIP English-based personality inventory to Portuguese to identify work-related personality traits while keeping it gender-neutral.

Theoretical Background

Personality traits influence a person’s way of thinking, feeling, and behaving (Hughes & Batey, 2017). These traits can be summarized using the highest level of descriptive behavior (Goldberg, 1993; McCrae & John, 1992) represented by the Big Five: Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience. Since personality is relatively stable (Costa et al., 2019), it can be used to predict people’s behaviors in their life outcomes (Zell & Lesick, 2022). The Big Five personality model (Goldberg, 1990) is a widely used and accepted taxonomy of traits (Costa et al., 1991) as it provides common ground for personality research and allows the classification of different human characteristics (Goldberg, 1990, 1993). The model has been studied in a wide range of industrial-organizational psychology topics,

linking personality with job outcomes (Goldberg, 1993), mostly with job performance (Barrick & Mount, 1991). Distinct job performance indicators have unique relationships with the Big Five factors. While Conscientiousness is the most robust predictor of job performance across occupations (Zell & Lesick, 2022), other factors have also been shown to be relevant depending on the nature of the job. Agreeableness is relevant to jobs that involve interpersonal relationships, and Openness to Experience is a valid predictor for customer service. Emotional Stability is significant with regard to customer service and teamwork, and Extraversion influences sales and managerial jobs (Barrick & Mount, 1991; Hurtz & Donovan, 2000).

Facets

Although the Big Five provides a useful framework for integrating and understanding human behavior at a higher level, it is difficult for the model to capture the whole diversity of human individuality (McCrae & John, 1992; Paunonen & Jackson, 2000). Therefore, narrow facets, a lower hierarchical conception of personality domains, are more useful to predict complex outcome variables such as job performance (Dudley et al., 2006; Hughes & Batey, 2017). Facets have shown greater predictive validity than their broader level factors (Dudley et al., 2006; Jenkins & Griffith, 2004; Lounsbury et al., 2003; Paunonen & Jackson, 2000; Seeboth & Möttus, 2018; Tett et al., 2003). Several personality models include a facet structure under the five broad factors: The NEO-PI-R (Costa & McCrae, 1995) contains 30 facets; the HEXACO model (Lee & Ashton, 2004) comprises 25 narrow facets; and the Hogan Personality Inventory (Hogan & Hogan, 2002), a commercial measure developed to predict occupational success, features 42 subscales.

Just as the definition of what constitutes successful job performance varies in different job roles and within different organizations, some facets may be more useful to predict specific work outcomes than others (Judge et al., 2013; Tett et al., 2003; Wilmot & Ones, 2021). In other words, facets yield distinct predictions depending on various performance criteria and the type of occupation in question (Dudley et al., 2006). For example, the facets Deliberation and Order from Conscientiousness were more related to task performance, while Dutifulness and Self-disciplined were more associated with interpersonal facilitation.

Context-Specificity in Personality Measures

The use of a contextual measure through the FOR effect (Bing et al., 2004; Holtrop et al., 2014; Schmit et al., 1995) has been shown to increase validity in the personality test by at least twice as much as noncontextualized measures (Shaffer & Postlethwaite, 2012). FOR occurs when a response to a personality scale differs considering the context

that the item behavior occurs. By specifying the context and using item tags (e.g., at school, at work, or at home), the individual focuses on how they would behave considering the context and with between-person variability reduced (Lievens et al., 2008, p. 2). For instance, the item *get upset easily* could be perceived differently by two individuals, depending on what personal information is most readily recalled at the time (Heller et al., 2007). One person may be concentrating on work-related matters, while the other may be concentrating on family-related concerns. When respondents were given an identical FOR, reliability increased and this was particularly noticeable at lower-level facets (Bing et al., 2014). The “at work” FOR has been shown to have a greater predictive power than general personality items (Hogan, 1991; Hunthausen et al., 2003; Lievens et al., 2008), predicting job satisfaction, job performance, work frustration, occupational stress, work engagement, career satisfaction, turnover intention, and absenteeism (Bowling & Burns, 2010; Burtaverde & Iliescu, 2019; Pace & Brannick, 2010). Furthermore, using the tag “at work” minimizes the need to “present oneself favorably” (Hunthausen et al., 2003). Job applicants are prone to adopting frames of reference outside of the work context when answering personality assessments lacking contextualization, as opposed to assessments that are contextualized (Fisher et al., 2017). They may also react negatively to the hiring process (Shaffer & Postlethwaite, 2012).

Culture, Language, and Personality Measures

The Big Five trait structure is considered universal (McCrae & Costa, 1997), indicating that its covariation pattern cuts across linguistic and cultural boundaries (Allik, 2005), including for the Portuguese population (Lima, 2002). Adaptations are necessary when we go deeper into the taxonomy, to the lower hierarchical levels, that might be impacted by culture. A few studies have raised issues when comparing trait scores across cultures (e.g., Geisinger, 1994), arguing that individuals from different cultures have differing reaction patterns, self-presentational rationales, comparison criteria, or even genetic foundations (Allik & McCrae, 2004). For instance, European and American cultures seem more extroverted and open to new experiences, while Asian, African, and Portuguese (although European) are more introverted, traditional, and conformist (Allik & McCrae, 2004). This is also confirmed through the lens of the Hofstede’s cultural dimensions (Hofstede, 2001). The Portuguese are known for their commitment to preserving strict moral and behavioral standards, often demonstrating a notable intolerance for unconventional ideas. Security plays a pivotal role in motivating individuals within this cultural framework. Furthermore, Portuguese culture is firmly rooted in

collectivism, with a strong emphasis on nurturing and cherishing interpersonal relationships. Cooperation is highly esteemed, which overshadows tendencies toward competition. Additionally, Portuguese society leans toward prioritizing tradition over long-term planning, and the prevailing ethos leans toward a culture of restraint rather than one that emphasizes enjoyment.

Adaptation implies more than mere translation and can even produce a completely new measurement (Van de Vijver & Leung, 1997). It should account for bias, such as construct bias (e.g., the construct measure is not identical to the original measure) and method bias (e.g., social desirability; Van De Vijver & Leung, 2001). Furthermore, adaptations must use gender-neutral language, as translating a questionnaire using gender-specific terms could hinder score interpretation and negatively impact the respondent's experience (Vainapel et al., 2015). Unlike English, Portuguese is marked by grammatical gender, specifically the reference to a person and the self-descriptions (e.g., PT *O acusado foi multado (masc.)/A acusada foi multada (fem.)* – EN *The accused was fined*; PT *Eu sou bonito (masc.)/eu sou bonita (fem.)* – EN *I am beautiful*). Since personality items tests are primarily self-descriptions, adaptation is necessary to keep it gender-neutral.

One example of a Portuguese IPIP adaptation is the Mini-IPIP five-factor model personality scale (Oliveira, 2019), which proved to have good psychometric properties and demonstrated that the five-dimensional structure was very reproducible in a Portuguese population. Another inventory that is an example of a Portuguese adaptation can be seen in the NEO-PI-R (Lima, 1997). While that particular instrument retained the five factors, there were variations in its thirty-facet structure that differentiated it from the American version. Additionally, eight out of the 240 items exhibited issues stemming from social desirability bias. Notably, both of these Portuguese scales were presented in the masculine form, which is the standard in Portuguese language usage.

This Research

Not only do several elements influence the predictive validity of personality tests, such as the frame-of-reference and narrow facets, but the personality of the Portuguese population differs from other Western cultures (Allik & McCrae, 2004) where the IPIP measures were originally developed. Furthermore, there is little agreement on what constitutes an optimal facet model (Costa & McCrae, 1998) and on the call to use narrow constructs from the Big Five to predict job outcomes (Hurtz & Donovan, 2000; Tett et al., 2003). Moreover, many of the existing inventories lack context-specificity,

which leads to uncertain results. Commercial publishers hardly ever allow extraction, changes, or rewording of items (Goldberg et al., 2006) and have shown low levels of reliability estimates ($< .70$). Considering these challenges, we set out to develop a work-related personality inventory containing a set of narrow facets structured under the Big Five domains. This measure is an adaptation of items from the IPIP (Goldberg et al., 2006), a public domain collection of items for use in personality tests. IPIP items are valuable predictors of job performance as long as a work-specific FOR is included (Shaffer & Postlethwaite, 2012). The inventory intends to assess work-related narrow personality traits, without referencing gender so that it can be used for personnel selection and for all individuals in the Portuguese working population. The study was structured into two phases: Study 1 involved the selection and adaptation of scales, item testing, and an analysis of scales interrelations. Study 2 centered on testing an enhanced instrument version and examining factor structures related to the Big Five personality traits. These analyses were carried out, taking into account the particularities of the labor context and the Portuguese language.

Study 1 – Pilot

Development of the Personality Inventory – Version 1

Selection of IPIP Scales

To consider and analyze more than 250 scales from the IPIP, two industrial organizational experts (one academic and the other professional – the author) selected the most relevant work scales. The selection was based on the following criteria: (1) constructs that are common in job postings (e.g., cooperation), (2) content of the scale items relevant to work (e.g., *Have an eye for detail*. vs. *Love children*), (3) both low and high scores of the scale are useful traits for a job (e.g., risk-taking may or may not be helpful for a job depending on the tasks required), and (4) have at least three facets for each Big Five factor (except for Emotional Stability). After the first selection, a second screening was conducted to compare item content similarity. As several scales had linguistic similarity, resulting in an overlapping construct, the scale with items more relevant to work was selected. This resulted in a total of 16 traits: Adaptability, Efficiency, Friendliness, Gregariousness, Leadership, Risk-taking, Self-efficacy, Inquisitiveness, Perfectionism, Cooperation, Creativity, Assertiveness, Empathy, Industriousness, Teamwork, and Emotional

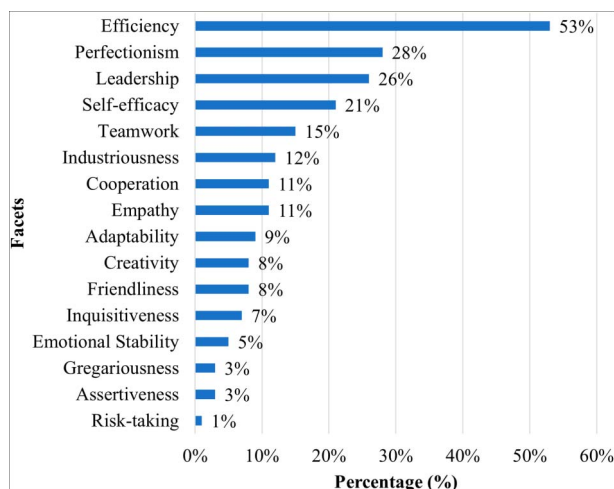


Figure 1. The percentage of facets in 429,239 job posts in English.

Stability. The decision to include Emotional Stability and not a lower-level facet was based on the scarce availability of scales that are relevant to work. Facets such as Anxiety and Fearfulness (Lee & Ashton, 2004) might confuse the respondent about the usefulness and pertinence of that information.

To check if these facets were suitable and applicable to current labor demands, we compared them with ongoing job openings. For each facet, we identified names or labels that might be deemed analogous to the facet (e.g., *leadership* – *people management*). Through web scraping, we took a sample of approximately 400,000 publicly available job descriptions. Using the spaCy natural language processing library (version 3.5.3), we employed a built-in named-entity recognition model, “en_core_web_md,” to compute the percentage of times each facet (along with their associated names) was requested in the job descriptions sample (see Figure 1). SpaCy is an open-source software library known for advanced natural language processing tasks (Honnibal & Montani, 2017). All selected facets showed applicability, even Risk-taking with 1%, which corresponds to 4,000 job descriptions.

Tagged Contextualization

Following the approach of previous frame-of-reference studies (Bing et al., 2014; Lievens et al., 2008), the tag “at work” was added to the beginning of each item. Just adding “at work” to existing IPIP items was inappropriate in some items, which necessitated rewording to maintain the connection to the construct (e.g., *At work I don’t know much about history* changed to *At work I try to understand the context of situations*; Lievens et al., 2008; Pace & Brannick, 2010; Schmit et al., 1995).

Portuguese Translation

The inventory was translated into Portuguese, followed by an independent back-translation (Brislin, 1980). Some items were not a literal translation but an adaptation, since some expressions were not commonly used in Portugal (e.g., *EN I am exacting in my work* was translated to PT *Tenho rigor no meu trabalho*).

Gender-Neutral Adaptation

The Portuguese-translated version contained 24 gender-referenced items. The items had trait adjectives that were by default translated to the masculine form. To keep the inventory gender-neutral, as in the English version, these items were rephrased: (1) from the first person (*I am*) to third person (*I am a person*; e.g., *I am a bad loser*; translation with gender PT *Sou um mau perdedor*; translation to third person PT *Sou uma pessoa com mau perder*), (2) from adjective to verb (e.g., *I can’t stand being contradicted*; translation with gender PT *Não suporto ser contrariado*; translation with verb PT *Não suporto que me contradigam*), and (3) from a male-adjective form to a two-gender adjective (e.g., *Am willing to try anything once*; translation with gender PT *Estou disposto para tentar qualquer coisa, pelo menos, uma vez*; translation with two-gender PT *Estou disponível para tentar qualquer coisa, pelo menos, uma vez*).

Methods

Participants

Through a nonprobability sampling technique, the snow-ball sampling effect, we collected 128 responses (1 participant was excluded due to unengaged responses). The sample comprised 127 participants, including 37 identifying as men, 86 identifying as women, and four participants who did not identify their gender. The age distribution was as follows: < 24 ($n = 2$), 25–30 ($n = 29$), 31–40 ($n = 58$), 41–50 ($n = 28$), 51–60 ($n = 2$), and > 60 ($n = 4$). Additionally, four participants chose not to disclose their age.

Measures

Participants were invited to complete an online questionnaire and asked to provide self-ratings for the items on a 5-point Likert-type rating scale, ranging from 1 = *strongly disagree* to 5 = *strongly agree*. Before the questionnaire, participants were required to read the following instructions: *You should take into account how you identify yourself in this moment at work, not how you want to be in the future. Assess yourself as you honestly see yourself*. The questionnaire included the following: (1) the work-related inventory – 135 items and (2) Goldberg’s markers for the Big Five (Goldberg, 1992) – 40 items from IPIP, 10 items for each domain (except for Emotional Stability) with the tag “at work.”

Statistical Analysis

We used two combined methods to identify unengaged and careless responses: (1) response time (Niessen et al., 2016) and (2) overuse of the same option (Johnson, 2005). For the first, we calculated the average time per item (6 s) and its *SD* (3 s). We set a cutoff of 9 min total response per questionnaire, around 3 s per item, since it was unlikely that the responses to all the items could be given in less than that (Meade & Craig, 2012). For the overuse of the same option, we counted the number of responses for each option and participant. We examined the resulting distribution to exclude participants with too many responses of the same kind.

Analyses were conducted using JASP 0.18. To maximize the power of the limited sample size ($N = 127$), we skipped the exploratory factor analysis and performed a confirmatory factor analysis (CFA) per facet. The goal was to extract factor loadings and understand how the items behave per facet. CFAs were calculated using DWLS (diagonally weighted least square) estimation. Model fit was assessed using the following goodness-of-fit indices: the chi-square test of model fit, the cumulative fit index (CFI), the root-mean-square of approximation (RMSEA), and the standardized root-mean-square residual (SRMR). Decisions to retain items were based on item content, factor loading, and goodness-of-fit indices. Reliability estimates were obtained using McDonald’s ω (McDonald, 1999) as the scale did not adhere to tau-equivalence which could be examined through the factor loadings (McNeish, 2018).

Results

Confirmatory Factor Analysis

The analysis resulted in general well-fitted models according to goodness-of-fit indices (Tables 1 and 2). The scales revealed a good model fit with a $p > .05$, a CFI $> .95$ indicating an adequate fit, and an RMSEA $< .06$ indicating a good fit – except for Empathy, Teamwork, Perfectionism, and Risk-taking.

Reliability Analysis

Reliability estimates obtained using McDonald’s ω can be found in Tables 1 and 2. Scale reliability coefficients can be grouped into four categories: Three scales had good coefficients ($\omega > .80$), seven scales had acceptable results ($\omega > .70$), five demonstrated poor consistency but were still acceptable ($\omega > .60$), and one was below .50.

Construct Validity

We examined the correlations among the scale scores of the facets to explore the relationship between them. Pearson correlations were determined using the scores from the facets (Table E1 in Electronic Supplementary Material 1 [ESM 1]). Some scales revealed high correlation ($r > .70$), indicating that they share more than 50% of similarities and suggesting possible construct overlap – such as Friendliness, Gregariousness, and Leadership; Assertiveness and Leadership; Efficiency and Industriousness; and Creativity and Inquisitiveness. Table E2 in ESM 1 provides correlations of the facets with Goldberg’s Big Five domains. Most of the

Table 1. Study 1 – CFA and reliability

| Item | Adaptability | Cooperation | Empathy | Teamwork | Efficiency | Industriousness | Perfectionism |
|----------------------|-------------------|-------------------|------------------|------------------|------------------|------------------|------------------|
| Item 1 | | .40 ^{ca} | .62 ^f | .83 | .50 | .42 | .75 |
| Item 2 | .26 ^a | .58 ^c | .46 | | .62 ^j | .42 | |
| Item 3 | .37 ^b | | .71 ^g | | .56 | | |
| Item 4 | .71 | .35 ^d | .59 ^h | .69 ⁱ | .53 | .74 | .75 |
| Item 5 | .66 | .91 | .71 ^h | | .56 | | |
| Item 6 | .89 | .41 ^e | .78 ^g | .39 | .52 | .72 ^k | .70 ^l |
| Item 7 | .26 ^{ab} | .66 | .63 ^f | .60 | .66 | .74 | |
| Item 8 | | .34 ^d | | .42 ⁱ | .36 ^j | .69 ^k | |
| Item 9 | | .52 | | | | | .76 ^l |
| Item 10 | | | | | | | .65 |
| Coefficient ω | .63 | .67 | .78 | .64 | .71 | .75 | .79 |
| $\chi^2(df)$ | 5.23(7) | 16.13(17) | 24.00(11) | 6.19(4) | 19.91(19) | 10.08(8) | 7.36(4) |
| p | .63 | .51 | .01 | .19 | .40 | .26 | .12 |
| CFI | 1.00 | 1.00 | .98 | .99 | 1.00 | .99 | .99 |
| RMSEA | .00 | .00 | .09 | .07 | .02 | .05 | .08 |
| SRMR | .03 | .04 | .06 | .04 | .05 | .04 | .03 |

Note. $N = 127$. For items, columns represent standardized loadings of the facet. Letters represent the residuals covariance between correspondent items: ^a=.25, ^b=.28, ^c=.31, ^d=.32, ^e=.20, ^f=.24, ^g=.21, ^h=.31, ⁱ=.26, ^j=.24, ^k=.18, ^l=.09.

Table 2. Study 1 – CFA and reliability

| Item | Assertiveness | Leadership | Friendliness | Gregariousness | Creativity | Inquisitiveness | Risk taking | Self efficacy | Emotional stability |
|----------------------|---------------|------------------|------------------|------------------|------------------|------------------|------------------|---------------|---------------------|
| Item 1 | .64 | .60 | .69 | .72 | | .56 | | .75 | .63 ^h |
| Item 2 | .63 | .52 ^a | .90 | .57 | .51 | .41 ^f | .54 ^g | .73 | |
| Item 3 | .66 | .44 ^a | .73 | .56 ^c | .41 ^d | .40 | .68 | | .65 ⁱ |
| Item 4 | .67 | .87 | .57 | .79 | .55 ^e | .54 ^f | | .60 | .84 |
| Item 5 | .76 | | .81 | .57 ^c | .54 ^e | .70 | .68 | | .71 |
| Item 6 | | | .63 | .45 | | | | .54 | .65 ^h |
| Item 7 | | | .58 ^b | .71 | .51 | | | | .85 ⁱ |
| Item 8 | | | .56 | .65 | .61 ^d | | | .76 | .59 |
| Item 9 | | | .78 ^b | | .91 | | .41 ^g | .62 | |
| Item 10 | | | | | | | | | |
| Coefficient ω | .78 | .64 | .87 | .81 | .75 | .58 | .60 | .77 | .86 |
| $\chi^2(df)$ | 4.93(5) | 0.56(1) | 27.07(26) | 25.23(19) | 14.55(12) | 5.30(4) | 1.8(1) | 4.91(9) | 12.27(12) |
| p | .42 | .45 | .41 | .15 | .27 | .26 | .18 | .84 | .42 |
| CFI | 1.00 | 1.00 | 1.00 | .99 | .99 | .99 | .99 | 1.00 | 1.00 |
| RMSEA | .00 | .00 | .02 | .05 | .04 | .05 | .08 | .00 | .01 |
| SRMR | .03 | .01 | .04 | .05 | .04 | .04 | .02 | .03 | .03 |

Note. $N = 127$. For items, columns represent standardized loadings of the facet. Letters represent the residuals covariance between correspondent items: ^a=.29, ^b=.26, ^c=.24, ^d=.30, ^e=.40, ^f=.28, ^g=.24, ^h=.31, ⁱ=.25.

facets were correlated with a domain; however, some showed correlations with multiple domains (e.g., Friendliness) and others had exceptionally strong correlations, such as Gregariousness and Extroversion ($r = .96$; $p < .001$), indicating an almost perfect positive linear relationship between the two constructs.

Discussion and Challenges in Translating

This first pilot study presented a few unexpected obstacles and the sample was also limited. In only three of the 16 scales, there were no issues regarding psychometric properties, such as internal consistency and construct validity. The remaining 13 were adapted and changed, as described in Study 2.

Upon comparing the performance of scale items that exhibited favorable results with those that performed poorly, we pinpointed specific issues: (1) poor wording resulted from the gender-neutral adaptation (e.g., Item 2 of Adaptability – *Am a bad loser*, the translation for Portuguese keeping the verb “to be” is PT *Sou uma pessoa com mau perder*; however, it should be translated using the verb “to have” to PT *Tenho mau perder*); (2) items were uncommon in professional situations (e.g., Item 6 of Cooperation; *Rarely overindulge*); (3) items were not clear or had a double meaning (e.g., Item 10 of Creativity; *Try to avoid complex people*); and (4) items had a different impact for the Portuguese; for instance, *Like to behave spontaneously* (Item

7 of Risk-Taking) may be seen as positive in Western cultures, but since Portuguese tend to be traditional and conformist (Allik & McCrae, 2004), this item was inappropriate for a work context. On the other hand, Perfectionism ($\omega = .79$), a scale that raised no issues, had items with simple translations, no double meaning or cultural influence, no adaptations to gender-neutral, and items were work-related (e.g., *Pay too little attention to details; Want everything to add up perfectly*).

Study 2

Development of the Personality Inventory – Version 2

After the analysis of the scales of Study 1, 72 items were deleted, 17 were rephrased, and 12 were added, leaving a total of 84 items. To address construct overlap issues, we proceeded to (1) merge the Efficiency and Industriousness scales; (2) remove Gregariousness and keep Friendliness, since the items were more appropriate for a work context; (3) reduce the Leadership items to focus on team management rather than items that relate to assertive behaviors (Posner & Kouzes, 1988); and (4) create new items for Creativity to differentiate from Inquisitiveness.

To increase internal consistency, some items were rephrased to become more evaluative or attractive

(Bäckström & Björklund, 2016) by aligning them with Portuguese cultural norms. The Adaptability scale was transformed into Receptiveness to enhance its alignment with the Agreeableness domain, emphasizing openness to others. Versatility was added to address being open to work changes and to have a construct associated with Openness to Experience. Since the scales Efficiency and Industriousness were merged, we added Organization to increase the number of facets that were associated with Conscientiousness. Cooperation was changed to Conformity as the items were related to accept standard norms. The inventory of version 2 resulted in 16 traits: Assertiveness, Conformity, Creativity, Emotional Stability, Empathy, Friendliness, Industriousness, Inquisitiveness, Leadership, Organization, Perfectionism, Receptiveness, Risk-taking, Self-efficacy, Teamwork, and Versatility.

Methods

Participants

A sample of 310 participants was collected through snowball sampling. The age distribution of study participants was as follows: < 24 years ($n = 19$), 25–30 years ($n = 46$), 31–40 years ($n = 99$), 41–50 years ($n = 53$), 51–60 years ($n = 25$), and > 60 years ($n = 11$), and 57 participants preferred not to answer their age. In terms of gender, there were 202 women, 51 men, and 57 participants who preferred not to disclose their gender. The dataset was randomly split into two subsamples with similar age and sex distributions. Each subsample had 155 participants.

Measures

Participants were invited to complete an online questionnaire and asked to provide self-ratings for the items on a 7-point Likert-type rating scale, ranging from 1 = *strongly disagree* to 7 = *strongly agree*. The questionnaire included (1) the work-related inventory – 84 adapted items and (2) Goldberg's markers for the Big Five (Goldberg, 1992) – 18 items from IPIP to measure the domains. The tag “at work” was added at the beginning of each item.

Statistical Analysis

The same method as Study 1 was used to identify unengaged or careless responses. Analyses were conducted using the same software as Study 1. We adopted a statistical approach similar to that employed by Rouco et al. (2022) in their study. Specifically, one CFA per facet was fitted using subsample 1. CFAs were conducted employing DWLS (diagonally weighted least squares) due to the presence of skewed response distributions in certain items, which were measured on ordinal scales. Model fit was

established based on the usual goodness-of-fit indicators. After we had calculated and summed up responses for the items' scales, subsample 2 was used to perform an EFA and determine the number of components underlying the facets to examine the factor structure. EFA was calculated via JASP 0.18 using varimax rotation and maximum likelihood (ML) estimation. Reliability estimates were obtained using McDonald's ω (McDonald, 1999).

Results

Confirmatory Factor Analysis

The goodness-of-fit indices showed that all measurement models fitted the data adequately. Tables 3 and 4 contain estimates of each facet's goodness of fit. However, we observed a notable presence of residual covariances among the items of the scales Empathy, Teamwork, Friendliness, and Versatility, indicating the existence of shared variance not explained by the facets.

Reliability Analysis

Reliability estimates were obtained using McDonald's ω for subsamples 1 (Tables 3 and 4) and 2 (see Table 5). Overall, the reliability coefficients were at least satisfactory for the majority of the facets ($\omega > .70$).

Structural Validity

Pearson correlations were calculated using the facet's scores (Table E3 in ESM 1). The maximum correlation obtained was between Self-efficacy and Creativity ($r = .61$), so there were no signs of complete construct overlap. EFA was performed to examine the factor structure of the facets. As seen in Table 5, only a 3-factor structure was found in the data, and three of the 16 scales showed cross-loadings. The retained factors (after rotation) accounted for 43% of variance. The first factor had an eigenvalue after rotation of 3.76, which was much larger than the second factor with 1.63, which explained only 10% of the variance. In addition, the scales were correlated with Goldberg's domains (Table E4 in ESM 1), and out of the 16 scales, 10 displayed correlations with Extraversion and Openness to Experience, and six of these correlations were observed with both of these Big Five domains.

Discussion

We intended to develop a gender-neutral and work-related personality inventory for the Portuguese population containing a set of narrow facets structured under the Big Five domains. Based on the item pool from the IPIP, we built an inventory with 16 scales showing good psychometric

Table 3. Study 2 – CFA and reliability

| Item | Empathy | Receptiveness | Teamwork | Conformity | Industriousness | Organization | Perfectionism | Assertiveness |
|----------------------|-------------------|------------------|-------------------|------------------|-----------------|------------------|---------------|-------------------|
| Item 1 | .86 | .89 | .78 ^{ef} | .52 ^g | .71 | .78 ^h | .70 | .69 ⁱ |
| Item 2 | .58 ^{ab} | .54 | .47 ^{df} | .81 | .51 | .79 | .71 | .54 ^{ij} |
| Item 3 | .64 ^b | .46 ^c | .53 ^{de} | .71 | .79 | .81 ^h | .74 | .66 ⁱ |
| Item 4 | .78 | .75 ^c | .84 | .22 ^g | .67 | .63 | .81 | .66 |
| Item 5 | .57 ^a | .67 | .69 | | .55 | | | .48 |
| Item 6 | | | | | | | | |
| Coefficient ω | .76 | .75 | .72 | .60 | .73 | .82 | .78 | .68 |
| $\chi^2(df)$ | 1.79(3) | 5.59(4) | 1.81(2) | 0.34(1) | 3.68(5) | 1.63(1) | 2.73(2) | 6.01(4) |
| p | .62 | .23 | .40 | .56 | .60 | .20 | .26 | .20 |
| CFI | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .99 |
| RMSEA | .00 | .05 | .00 | .00 | .00 | .06 | .05 | .06 |
| SRMR | .01 | .03 | .01 | .00 | .02 | .01 | .02 | .01 |

Note. Sample 1 ($n = 155$). For items, columns represent standardized loadings of the facet. Letters represent the residuals covariance between correspondent items: ^a=-.18, ^b=.17, ^c=.16, ^d=.47, ^e=-.32, ^f=.23, ^g=.18, ^h=-.18, ⁱ=.32, ^j=.13.

Table 4. Study 2 – CFA and reliability

| Item | Friendliness | Leadership | Creativity | Inquisitiveness | Risk-taking | Self-efficacy | Versatility | Emotional stability |
|----------------------|-------------------|------------|------------|------------------|------------------|------------------|-------------------|---------------------|
| Item 1 | .69 ^a | .57 | .79 | .56 | .72 ^d | .54 ^e | .78 | .76 |
| Item 2 | .66 ^{ab} | .43 | .67 | .43 ^c | .40 | .73 | .58 ^{fg} | .85 |
| Item 3 | .61 | .89 | .53 | .63 ^c | .50 | .45 | .58 ^g | .69 ^h |
| Item 4 | .82 ^b | .76 | .48 | .46 | .68 ^d | .44 | .44 ^f | .68 |
| Item 5 | .70 | | .81 | | | .57 ^e | .90 | .54 ^h |
| Item 6 | | | | | | .69 | .70 | |
| Coefficient ω | .75 | .76 | .76 | .54 | .78 | .67 | .73 | .78 |
| $\chi^2(df)$ | 0.55(3) | 3.8(2) | 5.5(5) | 0.94(1) | 0.63(1) | 13.87(8) | 2.38(7) | 3.64(4) |
| p | .91 | .15 | .36 | .33 | .43 | .09 | .94 | .46 |
| CFI | 1.00 | .99 | 1.00 | 1.00 | 1.00 | .98 | 1.00 | 1.00 |
| RMSEA | .00 | .08 | .02 | .00 | .00 | .07 | .00 | .00 |
| SRMR | .00 | .03 | .03 | .02 | .01 | .04 | .01 | .02 |

Note. Sample 1 ($n = 155$). For items, columns represent standardized loadings of the facet. Letters represent the residuals covariance between correspondent items: ^a=.17, ^b=.13, ^c=.13, ^d=-.34, ^e=.25, ^f=.29, ^g=.15, ^h=.23.

qualities. The reliability estimates for the scales had an average of .75 and all facets fitted the data according to goodness-of-fit indices.

The inventory was organized into three higher-order factors. In accordance with the Big Five model, factor 1 could be interpreted as a blend of Extroversion and Openness to Experience, factor 2 as a mix of Agreeableness and Emotional Stability, and factor 3 as Conscientiousness. Ashton et al. (2009) argue that a blended variable model, as opposed to the higher-order factor model, more clearly explains the relationships between the personality traits that make up the Big Five factors. One reason for these relationships is that traits such as creativity and leadership frequently exhibit high levels of both Extraversion and Openness. These characteristics determine someone's "social stimulus value" (Ashton et al.,

2009, p. 82) and their ability to attract the attention of others. In contrast, people who score low on these traits tend to be more conformist and passive observers. These qualities are less valued in the workplace and are rarely mentioned in job descriptions. As suggested by Ashton et al. (2009), it is not wise for personality inventories to possess a simple structure. Instead, they suggest that these inventories can better represent personality by looking at various aspects within each major personality dimension.

As a work-related tool, the 16 scales closely resemble Bartram's (2005) Great Eight Competencies. These are a set of core competencies used to provide a comprehensive framework for assessing an individual's potential and performance in the workplace. Bartram's principal component analyses of the predictor produced three-factor solutions accounting for 55% of the variance (Bartram, 2005).

Table 5. Study 2 – EFA and reliability

| Variable | Coefficient ω | Factor 1 | Factor 2 | Factor 3 |
|---------------------|----------------------|----------|----------|----------|
| Leadership | .78 | .76 | | |
| Self-efficacy | .69 | .71 | | |
| Risk-taking | .64 | .69 | | |
| Assertiveness | .70 | .69 | | |
| Creativity | .68 | .67 | | |
| Versatility | .78 | .60 | | |
| Inquisitiveness | .65 | .52 | | |
| Friendliness | .80 | .48 | .52 | |
| Industriousness | .76 | .45 | | .47 |
| Conformity | .82 | -.32 | | .44 |
| Receptiveness | .75 | | .59 | |
| Teamwork | .85 | | .51 | |
| Empathy | .73 | | .48 | |
| Emotional stability | .85 | | .45 | |
| Organization | .78 | | | .70 |
| Perfectionism | .73 | | | .66 |
| Eigenvalues | | 3.76 | 1.63 | 1.46 |
| Proportion var. | | .23 | .10 | .09 |
| Cumulative | | .23 | .34 | .43 |

Note. Sample 2 ($n = 155$). Applied rotation method is varimax. Columns represent standardized loadings of the facet. Eigenvalues after rotation.

Factor 1 comprised the competencies Leading/Deciding, Interacting/Presenting, Creating/Conceptualizing, and Enterprising/Performing. The scales comparable to our structure are Leadership, Assertiveness, Creativity, and Industriousness, respectively. Factor 2 included the competency Supporting/Cooperating, which is in line with the Teamwork and Empathy scales. As suggested by Bartram (2005), the focus on personality questionnaires, while valuable, has sometimes overshadowed the true importance of measurement in relation to workplace performance and outcomes.

While our study did not directly compare contextual and noncontextual scales to evaluate the effectiveness of FOR and increased reliability (Bing et al., 2014), we observed that our inventory, when compared with the Portuguese-adapted NEO-PI-R (Lima, 2002), exhibited a notably higher average reliability estimate (.75), surpassing NEOs (.56). Although FOR *at work* might influence social desirability, our instrument was applied in nonwork settings, alleviating this pressure. However, we believe that societal and cultural norms still exerted influence, a common challenge in our research context (Van De Vijver & Leung, 2001). Portuguese individuals often exhibit traits associated with reservedness and traditionalism (Allik, 2005; Lima, 2002). Items such as *I am relaxed most of the time* or *I am willing to try anything once* may have elicited less accurate responses

compared to *I am a very private person* or *I pay attention to details*, which yielded more favorable outcomes.

Implications

This inventory allows the collection of specific work-related traits, providing organizations with deeper personality data that can support their important decisions. It should also assist in the creation of a more substantial database for identifying the facets that are most pertinent to specific roles within corporations (He et al., 2019). The inventory represents personality at work more completely by assessing several distinct facets of each of the major personality dimensions. Additionally, in more recent years, there has been a discussion around diversity and inclusion addressing gender and inclusive pronouns; thus, it is critical to replicate these trends in assessment instruments, particularly those that are translated from English to a language with grammatical gender (e.g., Portuguese, French, German).

Limitations and Future Research

This research has some limitations that suggest important directions for future work. First, the sample of Study 1 was too small ($N = 127$), which prevented us from adopting an exploratory approach for a test that is being translated (Ziegler, 2020). Second, gender parity was not attained in the samples since men made up less than 30% of the overall sample in both studies. Furthermore, the impact of a gender-neutral version should be tested more thoroughly by comparing it to a gender-referenced inventory and observing the participants' reactions. Third, this study did not include noncontextual scales and, therefore, could not verify the increase in reliability due to work-specific scales (Bing et al., 2014). Fourth, the scales were selected from the IPIP, a pool of more than 250 scales, not an inventory, which resulted in an increased similarity between items and scales. We were not aware of a tool that can be used to identify redundancy among scales by detecting semantic overlap in psychological scales (Rosenbusch et al., 2020). Fifth, we recommend a study that compares these facets to others used in the IPIP-inventories (e.g., NEO-PI) to understand whether that inventory's facets have a better predictive validity than other inventories. Sixth, given the presence of not explained variance in some scales, alternative measurement models that better capture the relationships among scale items, should be explored. Ultimately, this inventory was designed for the evaluation of professional traits. Nevertheless, we did not investigate its external (or predictive) validity, which pertains to factors such as job performance or job engagement. To enhance the psychometric evidence, particularly in terms of predictive validity, we recommend considering the incorporation of computerized adaptive testing, as suggested by Qiu et al. (2022).

Conclusion

We have developed an assessment tool designed to evaluate 16 work-related traits within three personality domains linked to the Big Five framework. We hold the view that this fresh inventory will prove valuable for HR professionals and researchers keen on exploring the interplay between personality and work-related results. Despite the challenges faced, we remain optimistic that our present discoveries can steer future investigations and enrich our understanding of how these tools can be tailored and applied effectively.

The final version of this inventory will be available on the eikko website (eikko, n.d.).

References

- Allik, J. (2005). Personality dimensions across cultures. *Journal of Personality Disorders*, 19(3), 212–232. <https://doi.org/10.1521/pedi.2005.19.3.212>
- Allik, J., & McCrae, R. R. (2004). Toward a geography of personality traits: Patterns of profiles across 36 Cultures. *Journal of Cross-Cultural Psychology*, 35(1), 13–28. <https://doi.org/10.1177/0022022103260382>
- Ashton, M. C., Lee, K., Goldberg, L. R., & de Vries, R. E. (2009). Higher order factors of personality: Do they exist? *Personality and Social Psychology Review*, 13(2), 79–91. <https://doi.org/10.1177/1088868309338467>
- Bäckström, M., & Björklund, F. (2016). Is the general factor of personality based on evaluative responding? Experimental manipulation of item-popularity in personality inventories. *Personality and Individual Differences*, 96, 31–35. <https://doi.org/10.1016/j.paid.2016.02.058>
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44(1), 1–26. <https://doi.org/10.1111/j.1744-6570.1991.tb00688.x>
- Bartram, D. (2005). The great eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology*, 90(6), 1185–1203. <https://doi.org/10.1037/0021-9010.90.6.1185>
- Bing, M. N., Davison, H. K., & Smothers, J. (2014). Item-level Frame-of-reference Effects in personality testing: An investigation of incremental validity in an organizational setting. *International Journal of Selection and Assessment*, 22(2), 165–178. <https://doi.org/10.1111/ijsa.12066>
- Bing, M. N., Whanger, J. C., Davison, H. K., & VanHook, J. B. (2004). Incremental validity of the frame-of-reference effect in personality scale scores: A replication and extension. *The Journal of Applied Psychology*, 89(1), 150–157. <https://doi.org/10.1037/0021-9010.89.1.150>
- Bowling, N. A., & Burns, G. N. (2010). A comparison of work-specific and general personality measures as predictors of work and non-work criteria. *Personality and Individual Differences*, 49(2), 95–101. <https://doi.org/10.1016/j.paid.2010.03.009>
- Brislin, R. W. (1980). Translation and content analysis of oral and written material. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology: Methodology* (pp. 389–444). Allyn and Bacon.
- Burtaverde, V., & Iliescu, D. (2019). Emic vs etic frame of reference personality assessment in the prediction of work-related outcomes. *Career Development International*, 24(7), 686–701. <https://doi.org/10.1108/CDI-10-2018-0273>
- Costa, P., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the revised NEO personality inventory. *Journal of Personality Assessment*, 64(1), 21–50. https://doi.org/10.1207/s15327752jpa6401_2
- Costa, P., & McCrae, R. R. (1998). Trait theories of personality. In D. F. Barone, M. Hersen, & V. B. Hasselt (Eds.), *Advanced personality* (pp. 103–121). Plenum Press. https://doi.org/10.1007/978-1-4419-8580-4_5
- Costa, P., McCrae, R., & Dye, D. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO personality inventory. *Personality and Individual Differences*, 12(9), 887–898. [https://doi.org/10.1016/0191-8869\(91\)90177-D](https://doi.org/10.1016/0191-8869(91)90177-D)
- Costa, P., McCrae, R. R., & Löckenhoff, C. E. (2019). Personality across the life span. *Annual Review of Psychology*, 70(1), 423–448. <https://doi.org/10.1146/annurev-psych-010418-103244>
- Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, 132(4), 1593–1640. <https://doi.org/10.1093/qje/qjx022>
- Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology*, 91(1), 40–57. <https://doi.org/10.1037/0021-9010.91.1.40>
- eikko. (n.d.). *Personality test*. <https://eikko.ai/>
- Fisher, D. M., Cunningham, S., Kerr, A. J., & Allscheid, S. P. (2017). Contextualized personality measures in employee selection: Extending frame-of-reference research with job applicant samples. *International Journal of Selection and Assessment*, 25(1), 18–35. <https://doi.org/10.1111/ijsa.12156>
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6(4), 304–312. <https://doi.org/10.1037/1040-3590.6.4.304>
- 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>
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1), 26–42. <https://doi.org/10.1037/1040-3590.4.1.26>
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26–34. <https://doi.org/10.1037/0003-066X.48.1.26>
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1), 84–96. <https://doi.org/10.1016/j.jrp.2005.08.007>
- He, Y., Donnellan, M. B., & Mendoza, A. M. (2019). Five-factor personality domains and job performance: A second order meta-analysis. *Journal of Research in Personality*, 82, Article 103848. <https://doi.org/10.1016/j.jrp.2019.103848>
- Heller, D., Watson, D., Komar, J., Min, J.-A., & Perunovic, W. Q. E. (2007). Contextualized personality: Traditional and new assessment procedures. *Journal of Personality*, 75(6), 1229–1254. <https://doi.org/10.1111/j.1467-6494.2007.00474.x>
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage.
- Hogan, R., & Hogan, J. (2002). The Hogan personality inventory. In B. De Raad & M. Perugini (Eds.), *Big five assessment* (pp. 329–346). Hogrefe & Huber Publishers.

- Hogan, R. T. (1991). Personality and personality measurement. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology*, Vol. 2 (2nd ed., pp. 873–919). Consulting Psychologists Press.
- Holtrop, D., Born, M. Ph., de Vries, A., & de Vries, R. E. (2014). A matter of context: A comparison of two types of contextualized personality measures. *Personality and Individual Differences*, 68, 234–240. <https://doi.org/10.1016/j.paid.2014.04.029>
- Honnibal, M., & Montani, I. (2017). *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing* [Software package]. <https://doi.org/10.5281/zenodo.1212304>
- Hughes, D. J., & Batey, M. (2017). Using personality questionnaires for selection. In H. W. Goldstein, E. D. Pulakos, J. Passmore, & C. Semedo (Eds.), *The Wiley Blackwell handbook of the psychology of recruitment, selection and employee retention* (1st ed., pp. 151–181). Wiley. <https://doi.org/10.1002/9781118972472.ch8>
- Hunthausen, J., Truxillo, D., Bauer, T., & Hammer, L. (2003). A field study of frame-of-reference effects on personality test validity. *The Journal of Applied Psychology*, 88(3), 545–551. <https://doi.org/10.1037/0021-9010.88.3.545>
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, 85(6), 869–879. <https://doi.org/10.1037/0021-9010.85.6.869>
- Jenkins, M., & Griffith, R. (2004). Using personality constructs to predict performance: Narrow or broad Bandwidth. *Journal of Business and Psychology*, 19(2), 255–269. <https://doi.org/10.1007/s10869-004-0551-9>
- Johnson, J. A. (2005). Ascertaining the validity of individual protocols from Web-based personality inventories. *Journal of Research in Personality*, 39(1), 103–129. <https://doi.org/10.1016/j.jrp.2004.09.009>
- Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., & Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance: Integrating three organizing frameworks with two theoretical perspectives. *Journal of Applied Psychology*, 98(6), 875–925. <https://doi.org/10.1037/a0033901>
- Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate Behavioral Research*, 39(2), 329–358. https://doi.org/10.1207/s15327906mbr3902_8
- Lievens, F., De Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology*, 93(2), 268–279. <https://doi.org/10.1037/0021-9010.93.2.268>
- Lima, M. P. (1997). *NEO-PI-R: Contextos teóricos e psicométricos OCEAN ou iceberg? [EO-PI-R: Theoretical and psychometric contexts OCEAN or iceberg]* [Doctoral thesis]. Faculdade de Psicologia e de Ciências da Educação. <https://repositorio.ispa.pt/handle/10400.12/1647>
- Lima, M. P. d. (2002). Personality and culture: The Portuguese case. In R. R. McCrae & J. Allik (Eds.), *The Five-Factor model of personality across cultures* (pp. 249–260). Kluwer Academic/Plenum Publishers. https://doi.org/10.1007/978-1-4615-0763-5_12
- Lounsbury, J. W., Sundstrom, E., Loveland, J. M., & Gibson, L. W. (2003). Intelligence, 'Big Five' personality traits, and work drive as predictors of course grade. *Personality and Individual Differences*, 35(6), 1231–1239. [https://doi.org/10.1016/S0191-8869\(02\)00330-6](https://doi.org/10.1016/S0191-8869(02)00330-6)
- McCrae, R., & Costa, P. (1997). Personality trait structure as a human universal. *The American Psychologist*, 52(5), 509–516. <https://doi.org/10.1037/0003-066X.52.5.509>
- 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>
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Psychology Press.
- McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, 23(3), 412–433. <https://doi.org/10.1037/met0000144>
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437–455. <https://doi.org/10.1037/a0028085>
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, 60(3), 683–729. <https://doi.org/10.1111/j.1744-6570.2007.00089.x>
- Niessen, S., Meijer, R., & Tendeiro, J. (2016). Detecting careless respondents in web-based questionnaires: Which method to use? *Journal of Research in Personality*, 63, 1–11. <https://doi.org/10.1016/j.jrp.2016.04.010>
- Oliveira, J. (2019). Psychometric properties of the Portuguese version of the mini-IPIP five-factor model personality scale. *Current Psychology*, 38(2), 432–439. <https://doi.org/10.1007/s12144-017-9625-5>
- Pace, V. L., & Brannick, M. T. (2010). Improving prediction of work performance through frame-of-reference consistency: Empirical evidence using openness to experience. *International Journal of Selection and Assessment*, 18(2), 230–235. <https://doi.org/10.1111/j.1468-2389.2010.00506.x>
- Paunonen, S. V., & Jackson, D. N. (2000). What is beyond the Big Five? Plenty. *Journal of Personality*, 68(5), 821–835. <https://doi.org/10.1111/1467-6494.00117>
- Posner, B. Z., & Kouzes, J. M. (1988). Development and validation of the leadership practices inventory. *Educational and Psychological Measurement*, 48(2), 483–496. <https://doi.org/10.1177/0013164488482024>
- Qiu, X.-L., de la Torre, J., Ro, S., & Wang, W.-C. (2022). Computerized adaptive testing for ipsative tests with multidimensional pairwise-comparison items: Algorithm development and applications. *Applied Psychological Measurement*, 46(4), 255–272. <https://doi.org/10.1177/01466216221084209>
- Rosenbusch, H., Wanders, F., & Pit, I. L. (2020). The semantic scale network: An online tool to detect semantic overlap of psychological scales and prevent scale redundancies. *Psychological Methods*, 25(3), 380–392. <https://doi.org/10.1037/met0000244>
- Rouco, V., Cengia, A., Roberts, R., Kemper, C., & Ziegler, M. (2022). The Berlin Multi-Facet personality inventory: An IPIP-based measure of Big Five personality facets. *Psychological Test Adaptation and Development*, 3(1), 23–34. <https://doi.org/10.1027/2698-1866/a000021>
- Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology*, 80(5), 607–620. <https://doi.org/10.1037/0021-9010.80.5.607>
- Schulze, J., West, S. G., Freudenstein, J.-P., Schäpers, P., Mussel, P., Eid, M., & Krumm, S. (2021). Hidden framings and hidden asymmetries in the measurement of personality – A combined lens-model and frame-of-reference perspective. *Journal of Personality*, 89(2), 357–375. <https://doi.org/10.1111/jopy.12586>
- Seeboth, A., & Möttus, R. (2018). Successful explanations start with accurate descriptions: Questionnaire items as personality markers for more accurate predictions. *European Journal of Personality*, 32(3), 186–201. <https://doi.org/10.1002/per.2147>
- Shaffer, J. A., & Postlethwaite, B. E. (2012). A matter of context: A meta-analytic investigation of the relative validity of contextualized and noncontextualized personality measures. *Personnel Psychology*, 65(3), 445–493. <https://doi.org/10.1111/j.1744-6570.2012.01250.x>
- Tett, R. P., Steele, J. R., & Beauregard, R. S. (2003). Broad and narrow measures on both sides of the personality-job performance relationship. *Journal of Organizational Behavior*, 24(3), 335–356. <https://doi.org/10.1002/job.191>

- Vainapel, S., Shamir, O. Y., Tenenbaum, Y., & Gilam, G. (2015). The dark side of gendered language: The masculine-generic form as a cause for self-report bias. *Psychological Assessment*, 27(4), 1513–1519. <https://doi.org/10.1037/pas0000156>
- Van de Vijver, F. J., & Leung, K. (1997). Methods and data analysis of comparative research. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), *Handbook of cross-cultural psychology* (2nd ed., pp. 257–300). Allyn & Bacon.
- Van De Vijver, F. J., & Leung, K. (2001). Personality in cultural context: Methodological issues. *Journal of Personality*, 69(6), 1007–1031. <https://doi.org/10.1111/1467-6494.696173>
- Wilmot, M. P., & Ones, D. S. (2021). Occupational characteristics moderate personality–performance relations in major occupational groups. *Journal of Vocational Behavior*, 131, Article 103655. <https://doi.org/10.1016/j.jvb.2021.103655>
- 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>
- Ziegler, M. (2020). Psychological test adaptation and development – how papers are structured and why. *Psychological Test Adaptation and Development*, 1(1), 3–11. <https://doi.org/10.1027/2698-1866/a000002>

History

Received December 30, 2022

Revision received February 15, 2024

Accepted February 16, 2024

Published online April 24, 2024

Section: Methodological Topics in Assessment

Acknowledgments

The authors wish to thank Filipa Sobral for her expert advice and initial conceptualization and Miguel Baptista for his contribution on the development of this personality inventory.

Conflict of Interest

The first and second authors are a part of the company EIKKO – HR Technology Solutions that distributes the personality inventory discussed in this article.

Authorship

Andrea Ritter, writing – original draft, writing – review and editing, formal analysis. Aristides Ferreira, methodology, supervision. Hugo Esteves, data curation, software, funding acquisition. All authors approved the final version of the article.

Open Science

The data, code, and materials that support the findings of this study are available on request from the corresponding author, Andrea Ritter. The data are not publicly available due to privacy restrictions.

Andrea Ritter

EIKKO – HR Technology Solutions
Instituto Pedro Nunes
Rua Pedro Nunes
3030-199 Coimbra
Portugal
andrea.ritter@eikko.ai