



Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Comparison of the inter-item correlations of the Big Five Inventory-10 (BFI-10) between Western and non-Western contexts

Joonha Park ^{a,*}, Karlijn L. van den Broek ^{b,c}, Navjot Bhullar ^{d,e}, Charles Adedayo Ogunbode ^f, Julie Aitken Schermer ^g, Rouven Doran ^h, Rahkman Ardi ⁱ, Daniel Hanss ^j, Daniela Acquadro Maran ^k, Mai Albzour ^l, Sibebe D. Aquino ^m, Arin H. Ayanian ⁿ, Razieh Chegeni ^o, John Bosco Chika Chukwuorji ^p, Violeta Enea ^q, Elahe Ghanbarian ^r, Jihane Ghorayeb ^s, Feng Jiang ^t, Ojewumi Aderemi Kehinde ^u, Samuel Lins ^v, Michael J. Lomas ^w, Su Lu ^x, Tiago Marot ^m, Winfred Mbungu ^y, Ginés Navarro-Carrillo ^z, Charles Onyutha ^{aa}, Marc Eric S. Reyes ^{ab}, Katariina Salmela-Aro ^{ac}, Tomas Sollar ^{ad}, Hajra Tahir ^h, Chee-Seng Tan ^{ae}, Jorge Torres-Marín ^{af}, Takashi Tsubakita ^{ag}, Tatiana Volkodav ^{ah}, Anna Włodarczyk ^{ai}, Radha Yadav ^{aj}

^a School of Management, NUCB Business School, Japan

^b Copernicus Institute of Sustainable Development, Utrecht University, Netherlands

^c Research Centre for Environmental Economics, Heidelberg University, Germany

^d School of Psychology, University of New England, New South Wales, Australia

^e Discipline of Psychology, Edith Cowan University, Australia

^f School of Psychology, University of Nottingham, UK

^g Department of Management and Organisational Studies, Faculty of Social Science, University of Western Ontario, Canada

^h Department of Psychosocial Science, Faculty of Psychology, University of Bergen, Bergen, Norway

ⁱ Department of Psychology, Universitas Airlangga, Indonesia

^j Department of Social Sciences, Darmstadt University of Applied Sciences, Germany

^k Department of Psychology, University of Torino, Turin, Italy

^l Department of Social and Behavioral Science, Birzeit University, Palestine

^m Psychology Department, Pontifical Catholic University of Rio de Janeiro, Brazil

ⁿ The Institute for Interdisciplinary Research on Conflict and Violence, Bielefeld University, Germany

^o PROMENTA Research Center, Department of Psychology, University of Oslo, Oslo, Norway

^p Department of Psychology, University of Nigeria, Nigeria

^q Department of Psychology, Alexandru Ioan Cuza University of Iasi, Iasi, Romania

^r Department of Human Ecology, University of Alberta, Canada

^s Department of Psychology, Zayed University, Dubai, United Arab Emirates

^t Department of Human Resources and Organisational Behaviour, University of Greenwich, Beijing, China

^u Department of Psychology, Obafemi Awolowo University, Ife, Nigeria

^v Faculty of Psychology and Education Sciences, University of Porto, Portugal

^w School of Health and Society, University of Salford, UK

^x School of Applied Social Sciences, De Montfort University, UK

^y Department of Civil and Water Resources Engineering, School of Engineering and Technology, Sokoine University of Agriculture, Morogoro, Tanzania

^z Department of Psychology, University of Jaén, Spain

^{aa} Department of Civil and Environmental Engineering, Kyambogo University, Uganda

^{ab} Department of Psychology, College of Science, University of Santo Tomas, Manila, Philippines

^{ac} Department of Educational Sciences, University of Helsinki, Helsinki, Finland

^{ad} Department of Psychological Sciences, Faculty of Social Sciences and Healthcare, Constantine the Philosopher University in Nitra, Nitra, Slovakia

^{ae} Department of Psychology and Counselling, Faculty of Arts and Social Science, Universiti Tunku Abdul Rahman, Kampar, Malaysia

^{af} Department of Research Methods in Behavioural Sciences, University of Granada, Granada, Spain

^{ag} Faculty of Management, Nagoya University of Commerce and Business, Japan

^{ah} Faculty of Pedagogy, Psychology and Communicative Studies, Kuban State University, Krasnodar, Russia

^{ai} Escuela de Psicología, Facultad de Humanidades, Universidad Católica del Norte, Antofagasta, Chile

^{aj} School of Business and Management, Christ University, Bangalore, India

* Corresponding author at: School of Management, NUCB Business School, 1-3-1 Nishiki Naka-ku, Nagoya 460-0033, Japan.
E-mail address: jpark@nucba.ac.jp (J. Park).

<https://doi.org/10.1016/j.paid.2022.111751>

Received 21 March 2022; Received in revised form 22 May 2022; Accepted 23 May 2022

Available online 7 June 2022

0191-8869/© 2022 Elsevier Ltd. All rights reserved.

ARTICLE INFO

Keywords:

BFI-10
 Personality traits
 Inter-item correlations
 Omega coefficient
 Brief scales
 Western vs. non-Western countries

ABSTRACT

The Big Five Inventory-10 (BFI-10; Rammstedt & John, 2007) is one of many short versions of personality inventories that measure the Big Five trait dimensions. Short versions of scales often present methodological challenges as a trade-off for their convenience. Based on samples from 28 countries ($N = 10,560$), the current study investigated inter-item correlations estimated using Omega coefficients within each of the five personality characteristics measured by the BFI-10. Results showed that inter-item correlations were significantly lower, in the sample data from non-Western countries compared with the Western countries, for three of the five personality traits, specifically Conscientiousness, Extraversion, and Emotional Stability. Our findings indicate that the psychometric challenges exist across different cultures and traits. We offer recommendations when using short-item scales such as BFI-10 in survey research.

1. Introduction

Personality traits reflect individual differences in consistent patterns of thoughts, feelings, and actions (McCrae & Costa, 1990). The structure of those traits tends to converge into five dimensions: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (labelled as Emotional Stability in our study, representing the opposite pole of Neuroticism, consistent with Ludeke & Larsen, 2017). These five personality traits have been supported in multiple studies in a wide range of cultures, languages (McCrae et al., 1998) as well as with various methods (e.g., McCrae et al., 2005). Because of the important roles in explaining individual differences, personality traits are often included in national and international social and economic surveys, such as the International Social Survey Programme (ISSP), the German Socio-Economic Panel (SOEP), the UK Household Panel, World Values Survey (WVS), and the Household, Income and Labour Dynamics in Australia (HILDA) Survey (Rammstedt & Beierlein, 2014).

1.1. The Big Five Inventory-10 (BFI-10)

Brief measures assessing personality are helpful for researchers who are faced with limited assessment time (Gosling et al., 2003). The BFI-10 (Rammstedt & John, 2007) is one of such brief measures assessing personality characteristics. Each personality trait in the BFI-10 is represented by a pair of items designed to capture the essence of the specific trait without being redundant in content. Initially, this short measure was developed using data from two cultural groups (USA and Germany) in two languages (English and German) to provide an instrument for cross-cultural research. Consequently, this measure is often used in studies across various cultural contexts. However, it is essential to note that the evidence of the psychometric properties of this measure is mainly derived from culturally-biased samples, which poses some challenges, as explained below (Ludeke & Larsen, 2017).

1.2. Challenges of using brief measures

Despite the convenience, brief measures are often criticized for lacking psychometric quality, especially poor reliability issues (Gosling et al., 2003). This is true for personality scales including the BFI-10 (Ludeke & Larsen, 2017; Rammstedt & Beierlein, 2014). Ludeke and Larsen's (2017) analysis of the 6th WVS data from 25 national samples highlights two problems. First, the inter-item correlations (estimated with Pearson correlations) vary substantially across countries. The presence of low, and at times, negative inter-item correlations seems to be more apparent in countries categorized as Westernized, Educated, Industrialized, Rich, and Democratic (WEIRD; Henrich et al., 2010). Indeed, the two countries, USA and Germany, included in the original scale (BFI-10) development study, are classified as WEIRD (Rammstedt & John, 2007). Second, the issue with varying inter-item correlations is observed in some personality traits (e.g., Openness) and not with other personality characteristics (e.g., Conscientiousness) (Ludeke & Larsen, 2017).

To assess inter-item correlations, three reliability statistics (Cronbach's alpha, Pearson correlations, and Spearman-Brown coefficients) are generally used. Higher scores of the correlation coefficient reflect higher average inter-item covariance of a set of items (Raykov & Marcoulides, 2011). Of these three commonly used reliability statistics, Cronbach's alpha is the most frequently reported in the literature. This coefficient is computed based on the assumption that all items measure the same underlying variable and that "they are equally strongly associated to that underlying variable" (Peters, 2018, p. 59). However, the assumption that this index represents the same true score for all scale items, also known as tau-equivalence, is often violated (Catalán, 2019; Dunn et al., 2014; Graham, 2006; Revelle & Zinbarg, 2009; Sijtsma, 2009). Spearman-Brown correlation coefficients have been recommended for assessing the reliability of 2-item scales (Eisinga et al., 2013). However, it also fails to meet the assumption that scale items are equal in their strength of relationship with the underlying construct. To overcome these limitations, researchers are increasingly recommending the use of McDonald's Omega (McDonald, 1999) as a better index of internal consistency compared to other reliability indices (Flora, 2020; Hayes & Coutts, 2020; Peters, 2014; Rammstedt & Beierlein, 2014). Using McDonald's Omega can be beneficial, because this statistics is suitable even when the items are not equally related to the underlying construct (Catalán, 2019; Trizano-Hermosilla & Alvarado, 2016). This method is also valuable for evaluating short-item scales because "reliability estimated using Omega coefficient does not increase or decrease with the number of items in the scale" (Rammstedt & Beierlein, p. 214).

1.3. The current study

The present study examines the reliability of the BFI-10 by estimating inter-item correlations within each personality trait using Omega coefficients. Encouraged by numerous recommendations for the statistics as reviewed, we expect to offer more acceptable inter-item correlations and provide better insights into potential variance among personality traits across different cultural contexts.

The study analyzed the inter-item correlations of the BFI-10 scale between Western and non-Western countries as well as those across the five personality traits. The original dataset was part of a larger project on climate anxiety (see Ogunbode et al., 2019), which later included data from three additional countries (UAE, Palestine, and India). Whereas Ludeke and Larsen (2017) interpreted the cultural differences using the WEIRD vs. non-WEIRD framework, we analyzed data obtained from 28 countries categorized as Western versus non-Western as used in the original project (see Ogunbode et al., 2021). This is mainly because the current data are considered non-representative. For example, most of the samples in the present study were comprised of university students, making it difficult to control for the level of education in the interpretation of differences between Educated and non-Educated groups as per the WEIRD versus non-WEIRD categorization. Furthermore, literature on cultural differences in rating or response styles often considers cultural differences between the West and non-West (particularly, East Asian) cultural contexts (e.g., Hamamura et al., 2008). Therefore, we

Table 1
Summary of participating countries, group (Western/non-Western), and sample age (in years) and gender.

Country	Group	Mean _{Age}	SD _{Age}	n _{Male}	n _{Female}	N _{Total}
Australia	Western	35.55	11.03	81	230	311
Brazil	Western	30.56	12.77	76	204	280
Canada	Western	18.26	1.10	100	205	305
Chile	Western	23.96	9.26	151	219	370
China	Non-Western	21.3	2.85	75	186	261
Finland	Western	26.93	7.09	127	470	597
Germany	Western	23.43	5.67	75	169	244
Italy	Western	21.48	2.46	64	222	286
India	Non-Western	24.51	5.79	126	97	223
Indonesia	Non-Western	19.37	1.60	66	269	335
Iran	Non-Western	24.63	4.86	119	196	315
Japan	Non-Western	19.9	1.12	215	69	284
Malaysia	Non-Western	20.4	1.68	79	179	258
Netherlands	Western	24.15	6.01	139	268	407
Nigeria	Non-Western	25.1	7.95	265	294	559
Norway	Western	23.94	4.65	70	185	255
Pakistan	Non-Western	20.56	2.43	99	140	239
Palestine	Non-Western	20.49	2.96	91	231	322
Philippines	Non-Western	19.43	3.10	415	915	1330
Portugal	Western	32.2	13.90	64	190	254
Romania	Non-Western	21.3	1.59	132	263	395
Russia	Non-Western	25.08	8.43	88	373	461
Slovakia	Non-Western	20.98	1.40	52	204	256
Spain	Western	23.69	6.19	113	476	589
Tanzania	Non-Western	24.23	3.58	176	59	235
Uganda	Non-Western	26.07	5.87	526	116	642
United Arab Emirates (UAE)	Non-Western	21.12	2.98	14	164	178
United Kingdom (UK)	Western	20.2	3.99	38	331	369
Combined sample		23.32	7.11	3636	6924	10,560

concluded that the Western versus non-Western classification is more appropriate for understanding the cultural differences in inter-item correlations of the BFI-10 items.

We focused on the following questions. First, what are the profiles of inter-item correlations across the 28 samples at the country-level? Second, are there inter-item correlation differences between Western and non-Western countries? Specifically, does the BFI-10 exhibit lower internal consistency in the non-Western study samples than the Western samples? Third, do certain personality traits show poor inter-item correlations regardless of country-specific samples (e.g., Openness)?

2. Method

2.1. Participants and measures

A summary of participants' demographic information from each country included in the study is presented in Table 1. Participants completed the BFI-10 (Rammstedt & John, 2007) on a 5-point Likert scale assessing Openness, Conscientiousness, Extraversion, Agreeableness, and Emotional Stability, along with a range of other variables included in the larger project on climate anxiety (see Ogunbode et al., 2019).

2.2. Analysis

We conducted all analyses in RStudio 1.4.1717. Countries were categorized into Western versus non-Western groupings using country classification from Ogunbode et al. (2021). Omega coefficients of the two items, assessing the same personality trait, were computed using the omega function in the psych package (Revelle & Condon, 2019). Inter-item correlation mean scores for each country were then computed. We transformed these correlations using Fisher's transformation because they did not follow a normal distribution. Table 2 lists the Omega coefficients for each personality trait in each country. Omega coefficients range from 0 to 1, with values closer to 1 representing good reliability (Catalán, 2019). We also provide the inter-item correlations for the five personality traits estimated by Spearman-Brown correlation coefficients in Appendix A.

As a preliminary analysis, we visualize the correlations using Omega coefficients by country and traits (see Fig. 1), a method adopted from Ludeke and Larsen (2017) who used Pearson correlation coefficients. Fig. 2 shows the distribution of correlations between the five personality traits.

We also conducted a general linear mixed model with a random effect for Country, where the Country variable was nested within the Country-Group variable (Western/non-Western), to account for the dependence of data within countries. The dependent variable was the transformed correlation between the pairs of personality trait items for each country and the predictors included Country-Group and Trait. The interaction between Country-Group and Trait was included to assess if the differences in correlations across Western and non-Western countries differed across the personality traits (i.e., whether the reliability of the scale across country groups held for some traits, but not for others). To obtain the differences in inter-item correlations between Western and non-Western countries for each personality trait, we re-ran the regression analysis without the main-effect of Country-Group (Keppel & Wickens, 2004).

3. Results

The results of the generalized linear mixed model showed a significant main effect of Country-Group, $F(1, 28) = 33.30, p < .001$; the inter-item correlations between traits were higher in the Western countries ($M = 0.47, SD = 0.11$) than in the non-Western countries ($M = 0.29, SD = 0.17$). The results also indicated a significant main effect of Trait, $F(4, 112) = 34.61, p < .001$; the inter-item correlations differed across the personality traits. We also found a significant interaction effect between Country-Group and Trait, $F(4, 112) = 8.58, p < .001$; the differences in inter-item correlations varied across personality traits in the Western and non-Western countries. Overall, the multivariate model demonstrated a good fit: marginal $R^2_{(\text{fixed effect only})} = 0.57$, conditional $R^2_{(\text{fixed and random effects})} = 0.65$. Fig. 3 shows inter-item correlation means and error bars for each personality trait between the Western and non-Western countries.

The significant interaction effect suggested that the Country-Group variable was nested within the Trait variable. Hence, we re-ran the regression analysis without the main effect of Country-Group to obtain the simple effects for each trait for illustrative purposes. The results showed significant between-group differences for Conscientiousness: $\beta = 0.17, t(122.73) = 2.49, p = .01, 95\% \text{ CI } [0.03, 0.30]$, Extraversion: $\beta = 0.42, t(122.73) = 6.24, p < .001, 95\% \text{ CI } [0.29, 0.56]$, and Emotional stability: $\beta = 0.42, t(122.73) = 6.22, p < .001, 95\% \text{ CI } [0.29, 0.56]$. There was no statistically significant effects for Agreeableness: $\beta = 0.09, t(122.73) = 1.28, p = .20, 95\% \text{ CI } [-0.05, 0.22]$ and Openness: $\beta = 0.06, t(122.73) = 0.84, p = .40, 95\% \text{ CI } [-0.08, 0.19]$.

4. Discussion

The current investigation of inter-item correlations of the BFI-10 scale with Omega estimation method provides both challenges with

Table 2
Omega (ω) correlation coefficients for the BF-10 traits across 28 countries.

Country	Openness	Conscientiousness	Extraversion	Agreeableness	Emotional Stability
Australia	0.14	0.39	0.69	0.44	0.68
Brazil	0.43	0.31	0.73	0.00	0.68
Canada	0.08	0.43	0.60	0.23	0.48
Chile	0.30	0.37	0.55	0.01	0.61
China	0.36	0.45	0.51	0.25	0.61
Finland	0.43	0.57	0.70	0.11	0.71
Germany	0.42	0.57	0.70	0.07	0.54
Italy	0.43	0.31	0.73	0.28	0.76
India	0.53	0.02	0.01	0.14	0.33
Indonesia	0.09	0.57	0.75	0.00	0.50
Iran	0.05	0.33	0.38	0.12	0.30
Japan	0.24	0.37	0.53	0.11	0.61
Malaysia	0.33	0.29	0.46	0.20	0.54
Netherlands	0.34	0.43	0.59	0.35	0.71
Nigeria	0.29	0.17	0.11	0.13	0.02
Norway	0.47	0.48	0.77	0.27	0.71
Pakistan	0.27	0.05	0.21	0.17	0.11
Palestine	0.46	0.29	0.04	0.003	0.26
Philippines	0.20	0.30	0.29	0.27	0.55
Portugal	0.55	0.43	0.75	0.32	0.82
Romania	0.31	0.19	0.39	0.28	0.52
Russia	0.62	0.36	0.66	0.09	0.22
Slovakia	0.42	0.37	0.22	0.20	0.59
Spain	0.51	0.33	0.70	0.28	0.74
Tanzania	0.37	0.18	0.15	0.18	0.04
Uganda	0.06	0.34	0.31	0.04	0.09
UAE	0.30	0.31	0.59	0.002	0.60
UK	0.15	0.39	0.59	0.39	0.60

the low reliability of short-item scales and potential solutions to improve them. Compared inter-item correlations estimated with Spearman-Brown correlations, the overall results with Omega estimation showed improved reliabilities for each personality trait across national samples. Our study also demonstrated different patterns of inter-item correlations for the BFI-10 between the Western and non-Western countries. In the country-group comparisons, the inter-item correlations for three of the five personality traits (Conscientiousness, Extraversion, and Emotional Stability) were significantly higher in the Western countries than in the non-Western countries. Also, the inter-item correlations were not equivalent across the five personality traits. Consistent with Ludeke and Larsen (2017), Openness exhibited the lowest inter-item correlations, whereas Emotional Stability had the highest inter-item correlations in our multi-country data. The results for Extraversion and Emotional Stability with large group variance are similar to what was observed in Ludeke and Larsen (2017). These consistent patterns call for caution when adopting the two-item scales for these personality traits in cross-cultural contexts, especially in the non-Western countries. For example, in the Western country group, the inter-item correlations for Emotional Stability ranged from 0.48 to 0.82. However, inter-item correlations for the same personality trait observed in some of the non-Western countries (e.g., Nigeria, Tanzania, Uganda) were very low (all $\omega_s < 0.10$, see Table 2). Contrary to Ludeke and Larsen (2017), overall inter-item correlations for Openness in the current study were not poor (c.f., the results in Spearman-Brown correlation coefficients, Appendix A). All in all, our study findings based on Omega estimations suggested that Agreeableness exhibited the lowest reliability coefficients in both country groupings, whereas Extraversion and Emotional Stability were found to be the most reliable in the Western context.

4.1. Inter-item correlations of short-item scales across cultures

Although the scale reliability estimates is possibly improved with Omega coefficients, our findings demonstrate that poor inter-item correlations still exist in some countries for certain traits. Why do some personality traits (e.g., Agreeableness) show lower inter-item correlations? And why are the inter-item correlations likely to be lower in the

non-Western countries than the Western countries? The low inter-item correlations might be explained by the limited number of items used to assess a specific personality trait, as in the BFI-10. For this reason, McCrae and colleagues originally recommended including multiple items (i.e., eight items per facet in the NEO-PI-R) to increase reliability (McCrae et al., 1998). Second, translating adjectives describing personality characteristics across different languages could be more difficult as there is often not a single word to reflect the original word representing a personality characteristic as available in the English language. This might explain poorer inter-item correlations in the non-Western countries when compared to the Western countries. In these contexts, people could find it more challenging to describe themselves on a de-contextualized trait, because of their predominance of situation-based attributions (Fiske et al., 1998). Moreover, the composition of two items per trait, reflecting one positive adjective and one negative adjective, might be perplexing for people in cultures where participants show high response styles such as low extreme response and high mid-point response. Indeed, such response styles are reported to be more salient for respondents from non-Western countries (Chen et al., 1995; Church, 2010), which might explain the current patterns of inter-item correlation differences between the Western and non-Western countries.

4.2. Recommendations

Based on the current findings, we provide a few suggestions to consider when deciding upon the appropriate measures to use or when faced with the reliability issues with using brief measures. First, our study supports the idea that short scales need to be used with caution when the study is conducted with non-Western samples. Second, using Omega coefficient provides an improved reliability index for short-item scales. Third, to consider more likely response styles in some cultures (Chen et al., 1995; Church, 2010), if the use of a simplified measure is unavoidable in cross-cultural research including non-Western countries, it would be useful to consider providing positive adjectives only.

In addition, the current literature offers some recommendations. First, it is important to consider the purpose of the survey instrument. Lower levels of Cronbach's alphas are suggested to be acceptable for

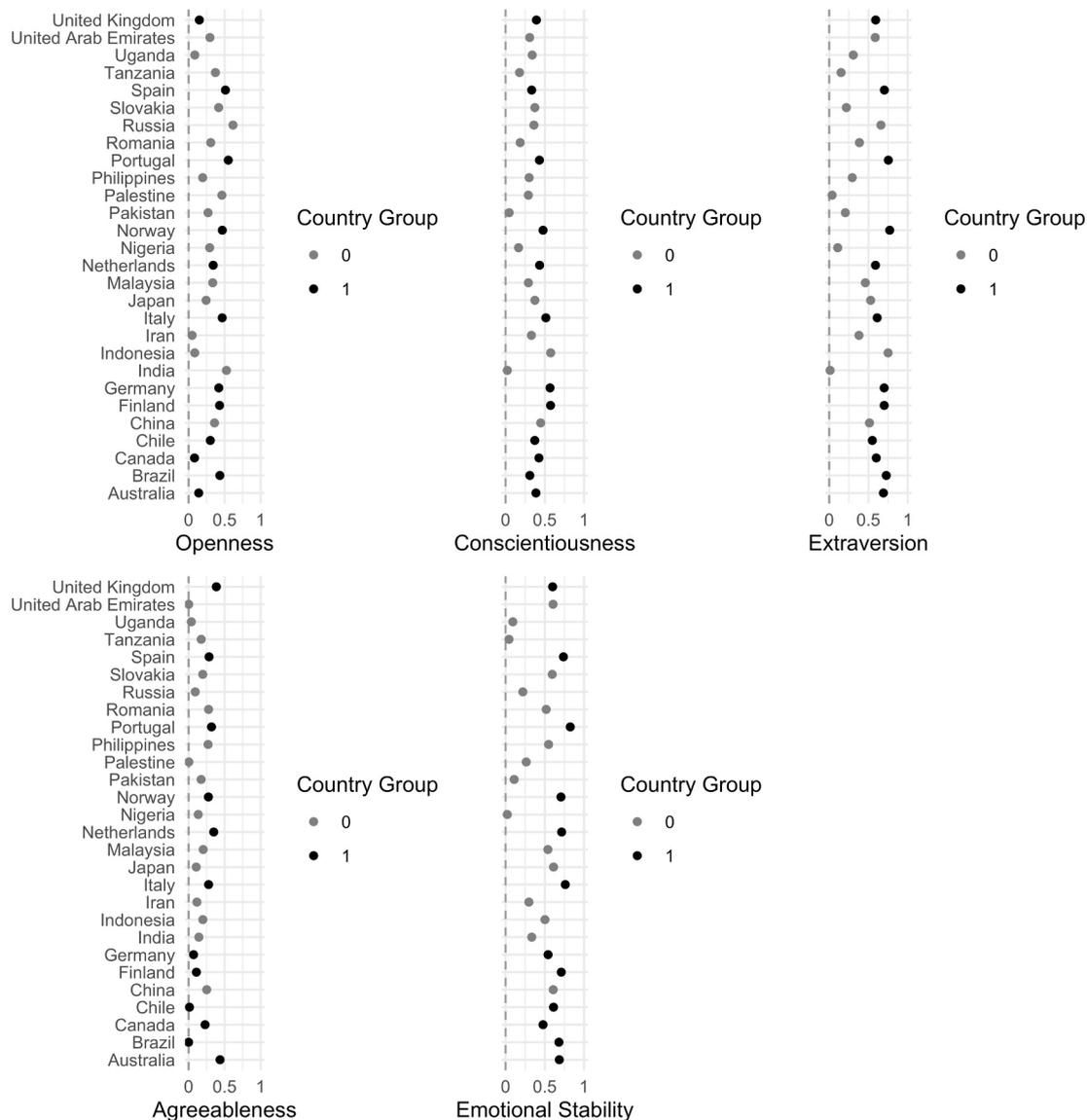


Fig. 1. Within-trait inter-item correlations coefficients (Omega) for the BFI-10 by country, grouped into Western (1) and non-Western (0) countries. Note. Adapted from Ludeke and Larsen (2017).

examining group-level differences, different from the inference of individual differences that requires higher levels of Cronbach's alphas (Nunnally & Bernstein, 1994). Thus, if the researchers are only interested in inferring group-level differences in psychological constructs, such short measures as the BFI-10 can be still an option, even in the non-Western regions. Fortunately, short measures seem to be usually adopted for large-scale, group-level investigations (Rammstedt & Beierlein, 2014).

Second, it is important for researchers to consider the opportunity cost between scale reliability and validity of a measure and clarify it beforehand based on their research purpose. It would also be helpful to consider the breadth of the psychological construct the researcher wants to measure. If the concept is generic, using short item measures would hardly show an adequate level of inter-item correlations. In the case of two-item scales, the broader the construct under investigation, the more extensive the range of the two items, resulting in lower inter-item correlations (Rammstedt & Beierlein, 2014). Indeed, the two items in each of the personality trait scales of the BFI-10 were developed to cover "as broad a bandwidth as possible for each scale, [so that both could capture] core aspects of the dimension but were not highly redundant in content" (Rammstedt & John, 2007, p. 205). Nonetheless, very low

inter-item correlations, and even negative correlations, as found in Ludeke and Larsen (2017), would question the unidimensionality of the trait.

Third, test-retest reliability is suggested to be an alternative way of examining the internal consistency (Gosling et al., 2003). Rammstedt and John (2007) estimated test-retest reliability of the BFI-10 in the two Western samples (USA and Germany) and found good levels of reliability in both countries, with mean retest stability coefficients being 0.72 and 0.78, respectively. It seems necessary to do the same in the non-Western country samples.

4.3. Conclusion

The current study, based on multi-country data, provides an alternative solution for the psychometric issue of using brief personality measures like the BFI-10 in cross-cultural studies (Gosling et al., 2003; Ludeke & Larsen, 2017). Overall, inter-item correlations tend to be lower than those measured with multiple-item scales. However, Omega coefficients can improve the issue of low inter-item correlations as observed with the short-item measures. Nevertheless, certain personality traits (e.g., Agreeableness) seem to exhibit lower reliability

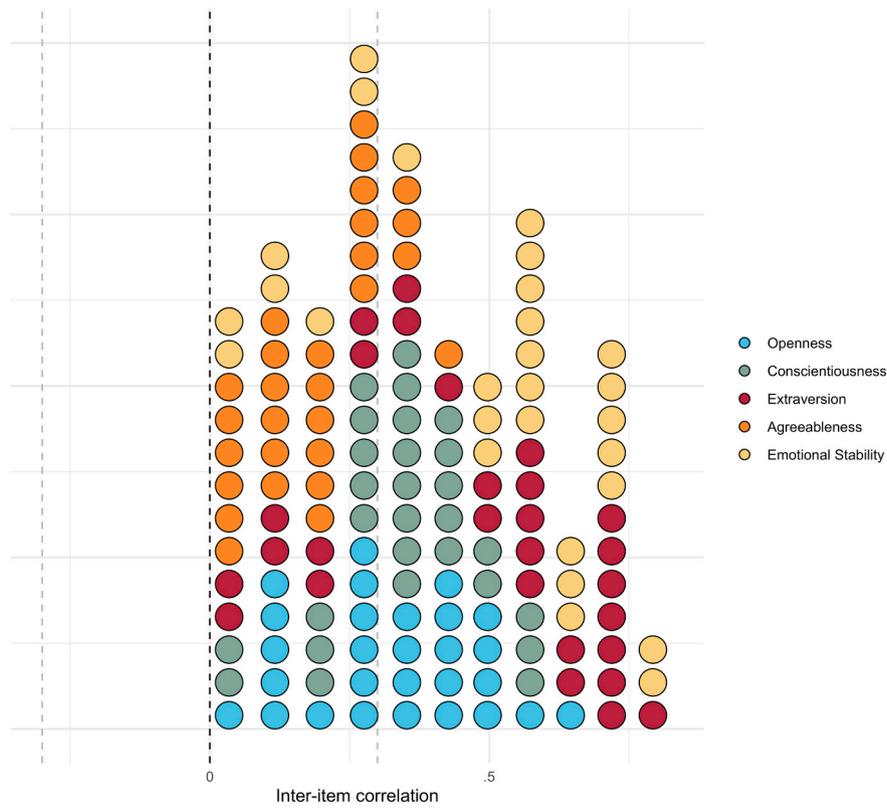


Fig. 2. Distribution of inter-item correlation coefficients (Omega) in the combined 28-country dataset. Note. Adapted from Ludeke and Larsen (2017).

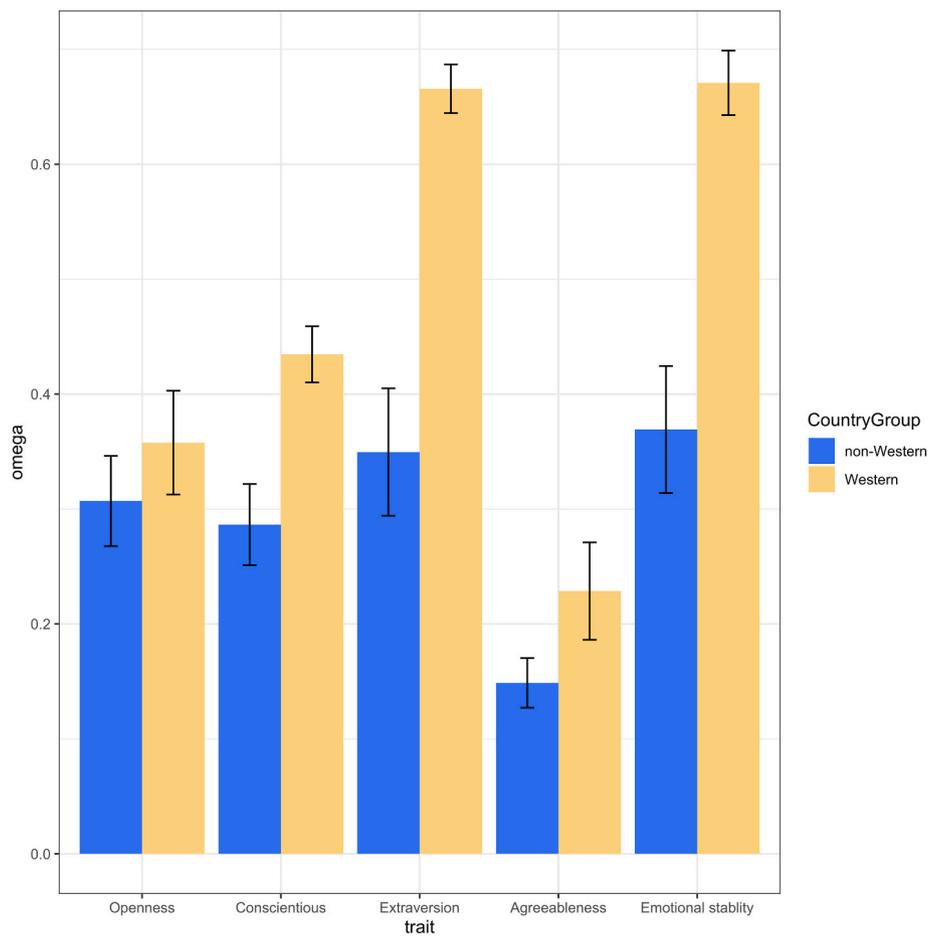


Fig. 3. Means and errors bars of inter-item correlation coefficients (Omega) for each personality trait by Western and non-Western country groups.

coefficients. There are also cultural differences in the attribution styles as well as response styles that might explain inter-item correlation differences observed in the Western vs. non-Western cultural contexts. Besides using Omega coefficients, we also recommend checking test-retest reliability as another index of internal consistency in the non-Western contexts as well as considering the breadth of the construct under investigation. We caution against using brief measures in the individual-level inference of the psychological construct. Our findings and recommendations offer insights to researchers using short-item scales such as BFI-10 in survey research in different cultural contexts.

CRedit authorship contribution statement

Joonha Park: Conceptualization, Data visualization, Data analysis, Writing- Original draft preparation, **Karlijn L. van den Broek:** Investigation, Writing – Results, Reviewing and Editing, **Navjot, Bhullar:** Investigation, Writing – Reviewing and Editing, **Charles Adedayo Ogunbode:** Original study design, Methodology, **Julie Aitken Schermer:** Investigation, Writing – Reviewing and Editing, **Rouven Doran:** Writing – Reviewing and Editing, **Rahkman Ardi:** Writing – Reviewing and Editing, **Daniel Hanss:** Writing – Reviewing and Editing, **Daniela Acquadro Maran:** Writing – Reviewing and Editing, **Sibele Aquino:** Data collection, **Arin H. Ayanian:** Data collection, **Razieh Chegeni:** Data collection, **JohnBosco Chika Chukwuorji:** Data collection, **Violeta Enea:** Data collection, **Elahe Ghanbarian:** Data collection, **Feng Jiang:** Data collection, **Samuel Lins:** Data collection, **Michael J. Lomas:** Data collection, **Su Lu:** Data collection, **Tiago Marot:** Data collection, **Winfred Mbungu:** Data collection, **Ginés Navarro-Carrillo:** Data collection, **Ojewumi Aderemi Kehinde:** Data collection, **Charles Onyutha:** Data collection, **Marc Eric S. Reyes:** Data collection, **Katariina Salmela-Aro:** Data collection, **Tomas Sollrar:** Data collection, **Hajra Tahir:** Data collection, **Chee-Seng Tan:** Data collection, **Jorge Torres-Marín:** Data collection, **Takashi Tsubakita:** Data collection, **Tatiana Volkodav:** Data collection, **Anna Wlodarczyk:** Data collection, **Radha Yadav:** Data collection.

Funding

The Chinese study has been supported by National Natural Science Foundation of China given to Feng Jiang [grant number: 71971225].

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.paid.2022.111751>.

References

Catalán, N. (2019). Reliability, population classification and weighting in multidimensional poverty measurement: A Monte Carlo study. *Social Indicators Research*, 142, 887–910. <https://doi.org/10.1007/s11205-018-1950-z>

Church, A. T. (2010). Measurement issues in cross-cultural research. In G. Walford, M. Viswanathan, & E. Tucker (Eds.), *The sage handbook of measurement* (pp. 151–177). Thousand Oaks, CA: Sage Publications.

Dunn, T. J., Baguley, T., & Brunsten, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399–412. <https://doi.org/10.1111/bjop.12046>

Eisinga, R., Grotenhuis, M. t., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health*, 58, 637–642. <https://doi.org/10.1007/s00038-012-0416-3>

Fiske, A. P., Kitayama, S., Markus, H. R., & Nisbett, R. E. (1998). The cultural matrix of social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 915–981). New York: McGraw-Hill.

Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science*, 484–501. <https://doi.org/10.1177/2515245920951747>

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 504–528. [https://doi.org/10.1016/S0092-6566\(03\)00046-1](https://doi.org/10.1016/S0092-6566(03)00046-1)

Hamamura, T., Heine, S. J., & Paulhus, D. L. (2008). Cultural differences in response styles: The role of dialectical thinking. *Personality and Individual Differences*, 44(4), 932–942. <https://doi.org/10.1016/j.paid.2007.10.03>

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But... *Communication Methods and Measures*, 14(1), 1–24. <https://doi.org/10.1080/19312458.2020.1718629>

Henrich, J., Heine, S., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33, 61–83. <https://doi.org/10.1017/S0140525X0999152x>

Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Saddle River, NJ: Pearson Prentice Hall.

Ludeke, S. G., & Larsen, E. G. (2017). Problems with the Big Five assessment. *Personality and Individual Differences*, 112, 103–105. <https://doi.org/10.1016/j.paid.2017.02.042>

McCrae, R. R., & Costa, P. T., Jr. (1990). *Personality in adulthood*. Guilford Press.

McCrae, R. R., Costa, P. T., Pillar, G. H., Rolland, J.-P., & Parker, W. D. (1998). Cross-cultural assessment of the Five-Factor model: The revised NEO personality inventory. *Journal of Cross-Cultural Psychology*, 29, 171–188. <https://doi.org/10.1177/0022022198291009>

McCrae, R. R., Terracciano, A., & 78 Members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, 88(3), 547–561. <https://doi.org/10.1037/0022-3514.88.3.547>

McDonald, R. P. (1999). *Test theory. A unified treatment*. Mahwah, NJ: Erlbaum.

Nunnally, J. C., & Bernstein, I. H. (1994). The assessment of reliability. *Psychometric Theory*, 3, 248–292.

Ogunbode, C. A., Doran, R., Hanss, D., Ytre-Arne, B., Moe, H., Ojala, M., & Salmela-Aro, K. (2019). Investigating associations between media exposure, climate anxiety and mental health (MECAMH). <https://osf.io/6n4rb/>.

Ogunbode, C. A., Pallesen, S., Böhm, G., et al. (2021). Negative emotions about climate change are related to insomnia symptoms and mental health: Cross-sectional evidence from 25 countries. *Current Psychology*. <https://doi.org/10.1007/s12144-021-01385-4>

Peters, G. Y. (2018, July 18). *The alpha and the omega of scale reliability and validity: Why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality*. <https://doi.org/10.31234/osf.io/h47fv>

Rammstedt, B., & Beierlein, C. (2014). Can't we make it any shorter? The limits of personality assessment and ways to overcome them. *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, 203–212. <https://doi.org/10.1016/j.jrp.2006.02.001>

Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. New York, NY: Taylor & Francis.

Revelle, W., & Condon, D. M. (2019). Reliability from α to ω : A tutorial. *Psychological Assessment*, 31(12), 1395–1411. <https://doi.org/10.1037/pas0000754>

Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: Congeneric and asymmetrical measurements. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00769>