Method Effects Associated to Item Valence: Evidence From The 10-item Big-Five Inventory In Older Adults

Irene Fernández[1], Zaira Torres[2]*, Sara Martínez-Gregorio[3], Amparo Oliver[4], & José M. Tomás[5]

Department of Methodology of the Behavioral Sciences, Faculty of Psychology, University of Valencia, Valencia, Spain.

[1] ORCID: 0000-0001-7333-3015

Professional degree: Master

[2] ORCID: 0000-0002-6344-6080

Professional degree: Master

[3]ORCID: 0000-0003-4274-4529

Professional degree: Master

[4] ORCID: 0000-0002-1207-4088

Professional degree: Doctoral degree, University professor

[5] ORCID: 0000-0002-3424-1668

Professional degree: Doctoral degree, University professor

Corresponding author: Zaira Torres

University of Valencia, Valencia, Spain, Department of Methodology for the Behavioral Sciences, Av. de Blasco Ibáñez, 21, 46010 Valencia, Valencia.

Phone number: 963864547

E-mail address: Zaira.Torres@uv.es

Method Effects Associated to Item Valence: Evidence From The 10-item Big-Five Inventory In Older Adults

Irene Fernández[1], Zaira Torres[2]*, Sara Martínez-Gregorio[3], Amparo Oliver[4], & José M. Tomás[5]

Department of Methodology of the Behavioral Sciences, Faculty of Psychology,

University of Valencia, Valencia, Spain.

Abstract

The objective of this study is to analyze the factor structure of the BFI-10 considering item valence effects when applied to measure older adults. Likewise, this study aims to estimate the factorial structure, internal consistency of the scale, to assess the nomological validity, and the association of the Big Five traits with age. 75078 participants with mean age of 68.27 from the 7th Wave of the SHARE study were included. Confirmatory Factor Analyses, omega coefficients and Pearson correlations were estimated. The best-fit model identified a five-factor structure with two valence effects, internal consistency ranged from .26 to .64, the nomological network showed that loneliness is negatively associated to neuroticism and positively with the other four traits, and the opposite direction in the associations with the five traits and satisfaction and quality of life. Consciousness, Extraversion and Openness have been found as dimensions that tend to decrease with age.

Introduction

Traditionally, the Big Five personality attributes model (namely extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience) has been one of the most used in the field of personality psychology (Thalmayer et al., 2011) and specially to describe personality throughout adulthood (Brandt et al., 2020). Personality traits are difficult to study because they are mainly constant throughout life, but they are also flexible, and people change as they age (Damian et al., 2019), go through the psychosocial changes of adulthood, and gain life experience (Costa et al., 2019). Generally, longitudinal personality studies in old adults have found a tendency for neuroticism to increase, and for other personality traits to decrease as people age (Berg and Johansson, 2014; Kandler et al., 2015; Mottus et al., 2012; Wagner et al., 2016), but others only support this decline as we grow older in Consciousness, Extraversion and Openness (Atherton et al., 2022). However, the Big Five personality traits showed relative stability in adulthood (Terracciano et al., 2006; Wagner et al., 2019), meaning that people are relatively stable on each personality variable over time or change in similar ways (Olaru & Allemand, 2020).

The gold standard for assessing the Big Five personality dimensions has been the NEO-PI-R (Costa & McCrae, 1992). Despite its great theoretical and psychometric properties, this instrument has the inconvenient of taking about 45 minutes to complete (Balgiu, 2018). Instruments with a reduced number of items avoid redundancy and long completion time (Guido et al., 2015). For this reason, researchers have developed alternative short personality instruments as the 60-item NEO-FFI (McCrae & Costa, 2004), the 44-item Big Five Inventory (John & Srivastava, 1999), Goldberg's (1992) set of 100 Trait Descriptive Adjectives (TDA), their compact version of 40 descriptors

(Saucier, 1994), or the 60-item BFI-2 scale (Soto & John, 2017). There are also extra-short measures to be used in large-scale surveys (Hofmans et al., 2008), such as the Ten-Item Personality Inventory (TIPI), the Five-Item Personality Inventory (FIPI, Gosling et al., 2003), the Single-Item Measures of Personality (SIMP, Woods & Hampson, 2005), and the one to discuss in this article, the Big Five Inventory-10 (BFI-10) designed by Rammstedt and John (2007).

The BFI-10 is a ten-item scale developed in two languages, English and German, with samples of students. It was thought to assess the Big Five dimensions of personality in a very short amount of time. The items of the scale were selected from the BFI-44, based on expert judgment and empirical item analyses. This scale represents an asset for health surveys, where resources and time limitations require administration of such short measures (Eisinga et al., 2013). In fact, in 2017 the BFI-10 was administered for the first time in the Survey of Health, Aging and Retirement in Europe (SHARE), a large, longitudinal, household survey of representative samples of European people aged 50 years and over (Börsch-Supan, 2020).

Regarding the factor structure of the BFI-10, multiple studies have examined whether the original five-factor structure was replicated, and mixed results have been found. In the first validation study, the expected five-factor structure of the BFI-10 was displayed in both US and German student samples (Rammstedt & John, 2007). While this original structure has been replicated in Italian (Guido et al., 2015) and French (Courtois et al., 2020) samples, item cross-loadings have also been documented, mostly for agreeableness items (Carciofo et al., 2016; Rammstedt et al., 2013) and to some extent for extraversion (Balgiu, 2018), neuroticism (Carciofo et al., 2016) and openness (Rammstedt et al., 2013) items. Finally, John et al. (2019) tested the five-factor model in a sample of Indian students but did not find convergence.

Overall, validation studies seem to disagree on the underlying structure of the BFI-10. There are two reasons that might obscure the scale's dimensionality. On the one hand, as noted by Kline (2016), problems in analysis are more common in factors with only two items. On the other hand, items can be framed in the direction of the construct under measurement, hence presenting a positive valence in respect to the construct under measurement, or in the reversed direction, with a negative valence. One way to create reversed items with negative valence is phrasing them with antonyms of that measured by the construct (Suárez-Álvarez, et al. 2018). Reversed items are thought to prevent response bias (Salazar-Solís, 2015). However, some authors have argued that reversing items poses interpretation problems that can lead to inattention and confusion (van Sonderen et al., 2013). Some other authors (Seng Kam & Meyer, 2015) have further demonstrated that social desirability response style partially explains method effects associated to item valence and item keying. In the case of the BFI-10 and item valence, there are only two items tapping each dimension, one framed positively and the other framed in the opposite direction of the construct. Hence, the little number of items per dimension and their item valence could account for divergent factor structures found among different validation studies. Furthermore, a key reason why the structure of personality is frequently challenging to replicate in representative samples of the adult population is that older people contain a wide range of cognitive abilities and different levels of education, closely related to method effects on the BFI-10 scale (Lechner and Rammstedt, 2015). For example, reverse-key items have been associated with several measurement problems (Menold, 2020) that could especially affect older adult samples, as they seem to require mental rotation of the rating scale to be processed, and this is a difficult cognitive task that depends on cognitive functioning (Kail and Park 1990). In

addition, poorer cognitive functioning has also been associated with higher levels of acquiescence (Rammstedt and Farmer, 2013).

Estimates of internal consistency of the BFI-10 show great disparity among studies, with reported Cronbach's alpha values ranging from .12 to .75 for Extraversion; .02-.78 for Agreeableness; .25-.46 for Conscientiousness; .33-.62 for Neuroticism; and .21-.76 for Openness (Carciofo et al, 2016; Hrebícková et al., 2016; John et al., 2019).

Although there are many studies that have tested the psychometric properties of the scale, most of them focus on samples of students or young adults, whereas the invariance of the measure across different age groups has been neglected or ignored in the personality scales (Dong & Dumas, 2020). A measure that does not show sufficient evidence of invariance, and is administered to a group different from its standardized population may lead to invalid personality scores. Therefore, there is a strong need to study the psychometric properties of this scale also in older people.

Regarding nomological validity, some authors have reported evidence on the association between the Big Five personality traits and several variables. Among the available measures in SHARE, some accumulate sufficient evidence of clear associations with personality, which means that they can be used to test for nomological validity. Specifically, the variables were: life satisfaction (Anglim et al., 2020; Balgiu, 2018; Lachmann et al., 2018; Rammstedt et al., 2020) and quality of life (Anglim et al., 2020; Balgiu, 2018), which, in turn, show negative associations with neuroticism and a positive relationship with extraversion, conscientiousness, agreeableness, and openness to experience. Loneliness is also an important variable in the context of old age, due to its higher prevalence, and its strong relation with personality traits (Hensley et al. 2012). Loneliness is positively related to neuroticism and negatively associated with the other

personality traits in a consistent way in older people (Itzick et al., 2020; Schutter et al., 2020; Wang & Dong, 2018).

Despite the amount of BFI-10 psychometric studies, there is a need to study the scale properties in samples of old people, and especially considering the potential method effects associated to item valence. The aim of this study is to test the psychometric properties of the BFI-10 in a large representative sample of European and Israelis adults. More concretely, the aims of the present study are a) to analyze potential item valence effects and establish the factor structure of the scale; b) to estimate internal consistency of the best-fitting model; c) to assess the nomological validity of the personality dimensions; and d) to test the association of the five personality dimensions with age.

Method

Sample and procedure

Data used in this study comes from the Survey of Health, Aging and Retirement in Europe (SHARE; Börsch-Supan *et al.*, 2013) on its 7th Wave (Börsch-Supan, 2020). The data are available at the SHARE Research Data Center to the entire research community free of charge (www.share-project.org) upon reasonable request and with permission of the SHARE project. For this study, the total sample was comprised by 77263 participants, from which 2185 were not administered the BFI-10. Hence, the final sample consisted of the remaining 75078 participants in this study.

From the final sample, 43041 (57.3%) participants were women and 32037 (42.7%) were men. Most of them were married and living together with their spouse (68.9%), while 15.2% were widows/widowers, 8.1% were divorced, and the remaining 7.8% were in other situations. Age ranged between 22 and 105 years old ($M = 68.27$, $SD = 9.90$). The amount of individuals younger than 50 years old accounted for the 1.1% of

the sample. Although the SHARE target population is all people aged 50 years and over, partners living in the same household are also eligible respondents regardless of their age. The sampling strategy employed in SHARE is a probabilistic four-stage process. For more details on sample eligibility and the sampling procedures, please see Bergmann *et al.* (2019). The Ethical Approval for gathering of the data used in this study was obtained by the SHARE project and are guided by the ethical declaration of Helsinki.

Instruments

Personality traits were measured using the 10-item Big-Five Inventory (BFI-10; Rammstedt & John, 2007). This measure of personality contains two items tapping each of the following factors: neuroticism, extraversion, conscientiousness, agreeableness and openness to experience. For each factor, one item is phrased in the direction of the construct and the other item is phrased in the opposite direction. Responses were coded in a 5-point Likert scale ranging from 1 (disagree strongly) to 5 (agree strongly).

Loneliness was measured with the Three-Item Loneliness Scale (Hughes et al., 2004), a reduced version of the R-UCLA Loneliness Scale (Rusell et al., 1980). The items of the scale assess the frequency of feelings of exclusion, isolation and abandonment. Responses are coded in a three-point Likert scale: 1 (hardly ever or never), 2 (some of the time), and 3 (often). Estimate of internal consistency was .76 by means of Cronbach's alpha.

Quality of life was assessed using the Control, Autonomy, Self-realization and Pleasure-12 (CASP-12) scale, a reduced version of the CASP-19 (Hyde et al., 2003). Item responses are coded in a four-point Likert scale, ranging from 1 (never) to 4 (often). The 12-item version of the instrument was recently validated using data from SHARE Wave 7 (Oliver et al., 2021). Results from the validation study support its use

as an overall score of subjective well-being. Scale's alpha was .83 for the general factor of well-being.

Life satisfaction was measured with the question: "On a scale from 0 to 10 where 0 means completely dissatisfied and 10 means completely satisfied, how satisfied are you with your life?".

Statistical analyses

To satisfy the first objective of the study, a set of Confirmatory Factor Analyses (CFAs) were estimated. Specifically, the four a priori models were: Model 1, a five-factor structure of the BFI-10 as established by Rammstedt & John (2007); Model 2, the five factor substantive structure plus a method effect factor associated to the five negative valence items of the scale; Model 3, the five factors of personality plus a method factor associated to the five positive valence items of the scale; and Model 4, with the five substantive factors plus both method effect factors, one associated to positive valence items and the other associated to negative valence items.

Assessment of model fit employed the recommended indices (Tanaka, 1993): chi-square statistic (χ^2), Comparative Fit Index (CFI), Root Mean Squared Error of Approximation (RMSEA) and Standardized Root Mean square Residual (SRMR). Adequate fit is considered with CFI equals or is over .90 and RMSEA/SRMR equal or are under .08, while excellent fit is deemed with CFI of at least .95 and RMSEA/SRMR of .05 or less (Hu & Bentler,1999). Our data were markedly non-normal and ordinal (five response alternatives), and therefore Weighted Least Squares Mean and Variance adjusted (WLSMV) was chosen as the method of estimation, given that this is the preferred method for compensating deviations from multivariate normality of ordinal data (Kline, 2016).

The second aim of the study was to estimate the internal consistency of the scale. For this, coefficient omega (ω, McDonald, 1999) was used, given that it overcomes the deficiencies of Cronbach's alpha. Specifically, it is known that coefficient alpha underestimates the true reliability unless the items are tau-equivalent, whereas coefficient omega does not (Deng & Chan, 2017). The third research objective was approached examining the relationships between personality dimensions, measured using the BFI-10, and loneliness, life satisfaction and quality of life, using Pearson's correlations. Finally, the fourth research objective was to analyze the association between age and the five personality traits using Pearson's correlations. Analyses were done with SPSS 26 and Mplus 8.7 (Muthén & Muthén, 1998-2021).

Results

Factorial validity

Fit indices of the four tested models are displayed in Table 1. The model that showed the best fit to the data was Model 4, which was hence retained as the best-fitting model. This model considered the five dimensions of personality together with two method factors, one associated to items framed in the direction of the measured dimensions and one associated to items framed in the opposite direction.

INSERT ABOUT HERE TABLE 1

Standardized factor loadings of the best-fitting model, Model 4, can be consulted in Figure 1. Both method factors were uncorrelated with each other and with the personality dimensions.

INSERT ABOUT HERE FIGURE 1

Additionally, it is a relevant information to calculate the personality dimensions' correlations when valence effects are taken into account and compare them with a model

in which valence effects are not controlled for (five-factor model without method effects). Table 2 shows this comparison, with a clear pattern occurring. Positive correlations among personality factors decreased in magnitude when valence effects were modelled, but negative correlations increased in magnitude. In one particular case, the correlation between agreeableness and openness, the correlation is negative when valence effects are controlled for, and positive when they are not.

INSERT ABOUT HERE TABLE 2

Internal consistency

Internal consistency was estimated using omega. Results were $\omega = .35$ for Openness, $\omega = .56$ for Consciousness, $\omega = .50$ for Extraversion, $\omega = .26$ for Agreeableness, and $\omega = .64$ for Neuroticism.

Criterion-related validity

Nomological validity was assessed with the correlations between the five dimensions of personality with loneliness, well-being and life satisfaction. Correlation coefficients and their associated *p*-values are displayed in Table 3. These (latent) correlation coefficients have been calculated with a CFA model with method effects and also with another CFA without method effects. This means that differences between these correlations inform about the effect it has to consider method effects on the correlation between the five personality factors.

INSERT ABOUT HERE TABLE 3

Correlations between the Big Five traits and age

The dimensions Extraversion r=-.043, (p<.001), Conscientiousness r= -.015, (p<.001), Openness r=-.066, (p<.001) and Neuroticism r=-.010, (p<.05) showed a decrease with increasing age, while Agreeableness increases r=.046, (p<.001) as people get older.

However, although significant all correlations are extremely low, thus indicating a great stability of personality trough the old age.

Discussion

The present study examined the psychometric properties of the BFI-10 (Rammstedt & John, 2007) using data from SHARE. Three psychometric properties were addressed. Firstly, evidence for its structure was examined testing a series of CFAs, in which special attention was paid to potential method effects associated to item valence. Secondly, internal consistency of the resulting factors was estimated using omega coefficients. Thirdly, criterion-related validity was studied using life satisfaction, loneliness and quality of life as criteria. Fourthly, the associations between age and personality traits were evaluated to see if they were similar to those found in previous studies.

Regarding factor validity, among the four tested models, the best-fitting model was the one including two method effects associated to positive and negative item valence. This study represents the first one to acknowledge such valence effects within the five-factor structure of the BFI-10. Some of the previous literature had found the simple five-factor structure of the Big-Five personality traits to fit the data (Courtois et al., 2020; Guido et al., 2015; Rammstedt & John, 2007), but some other studies had found some kind of misbehavior with item cross-loadings (Balgiu, 2018; Carciofo et al., 2016; Rammstedt et al., 2013). Even one study was not able to find model convergence for the five-factor structure (John et al., 2019). Nevertheless, some of these investigations consider that possible problems in the factorial structure may be due to the translation of the items, the item interpretation, or the cultural differences in personality structure (Carciofo et al., 2016). This should be considered especially in non-Western countries, as these cultural differences in personality may be more marked (John et al., 2019).

About correlations among the personality traits of the best-fitting model, they are generally as expected and reported in the literature on the Big-Five factors of personality, except for the one between agreeableness and openness to experience. In our study, these two dimensions correlated negatively. This result differs from previous studies, which found near-zero estimates for this relationship (Carciofo et al., 2016; Rammstedt & John, 2007) at the observable level, while John et al. (2019) found a positive association.

One possibility is that valence effects were obscuring the scale's dimensionality, as it is well known that mixing items with positive and negative wording creates problems in interpreting item content (van Sonderen et al., 2013). However, it could also be that just two items per dimension do not suffice to fully represent the construct (Marsh et al., 1998) or to provide an adequate analytical solution. As stated by Kline (2016), convergence and estimation problems are more frequently encountered in factors with few items.

In this line, internal consistency problems have also been documented, as it is difficult to obtain consistent responses to items belonging to the same construct when these are scant (Eisinga et al., 2013). Hence, one possible reason why so much disparity in internal consistency of the BFI-10 has been found. In the present study, internal consistency estimates using coefficient omega were low, as in the case of all previous studies looking at the scale's reliability either by means of Cronbach's alpha (Balgiu, 2018; Carciofo et al., 2016; John et al., 2019), Spearman-Brown coefficient (Guido et al., 2015; Hrebícková et al., 2016), or test-retest (Courtois et al., 2020). When considering these low estimates of internal consistency, we must be aware of the natural tension between internal consistency and content validity almost inherent to short scales that try to tap constructs with different facets. Indeed, as acknowledged by Little et al. (1999), accurate relationships among constructs can occur even with indicators with poor

reliability, given that they are well-selected across the construct's domain, variability is ensured and confirmatory factor analysis supports such factor structure.

Regarding nomological validity, results from this study replicated those by previous studies. Loneliness was positively associated to neuroticism and negatively related to the other four personality traits. As in the case of all previous studies looking at this relationship in old age samples (Itzick et al., 2020; Schutter et al., 2020; Wang & Dong, 2018), the relationship of loneliness with neuroticism was stronger than the relationship with the other personality dimensions. Life satisfaction and quality of life, on the other side, showed negative associations to neuroticism and positive ones to extraversion, agreeableness, conscientiousness and openness. This goes in line with previous literature (Anglim et al., 2020; Balgiu, 2018; Lachmann et al., 2018; Rammstedt et al., 2020). In general, as in the aforementioned studies, neuroticism showed the strongest association to the criteria, and openness the weakest. Hence, it seems that results found in this article suggest the BFI-10 to display adequate criterion-related validity.

Finally, the relationship found between the Big Five and age was as expected for the dimensions Extraversion, Conscientiousness and Openness. However, an increase in agreeableness and a decrease in neuroticism were found with increasing age. A possible explanation is that the sample also includes people in middle adulthood and during these years, agreeableness, tends to increase and neuroticism to decrease due to maturation and increased responsibilities (Marsh et al., 2013). In any case, these correlations have been relatively small, showing that personality traits are fundamentally stable.

All in all, results show that responses to the BFI-10 seem to interfere with item valence, and there is a lack of internal consistency, probably due to the little number of items per dimension. Seng Kam and Meyer (2015) argue that item valence ought to be controlled for using Multi-Trait Multi-Method (MTMM) CFA models, while deletion of

negative valence items is not recommended as these authors demonstrated that item valence influences nomological network analyses. Be that as it may, once item valence is acknowledged within the factor model, nomological validity can be established. Additionally, the use of short measures is recurrent in large scale surveys in which resources and time are limited (Eisinga et al., 2013). SHARE represents one of the most complete surveys in Europe in terms of sample representativeness, types of measures covered (social, psychological, economic, physical, etc.) and quality of research design. Despite the use of short measures not representing the best assessment scenario, balance between quality and scope of research ought to be contemplated. In this sense, SHARE provides an excellent middle ground of the quality of assessment instruments and methodology.

Among the contributions of this research, this is the first validation study of the BFI-10 in a representative sample of adults, which allows for application of the instrument beyond student populations. Moreover, this study expands the knowledge of valence effects in the BFI-10 scale and emphasizes on considering them when used in samples of older people, as our data show that method effects are present and are not negligible. Due to the importance of these effects, alternatively, it should be consider the possibility to incorporate constrained Factor Mixture Analysis (FMA) as proposed by Arias et al. (2020) and Steinmann et al. (2021) to screen out inconsistent respondents to mixed-worded scales. Findings should be, however, interpreted considering some limitations. Our study does not include concurrent and discriminant validity analyses, as the SHARE project only includes the BFI-10 personality measure. Although the BFI-10 scale has been correlated with other personality instruments in previous studies (Carciofo et al., 2016; Courtois et al., 2020; Guido et al., 2015; Rammstedt & John, 2007) and has shown evidence of concurrent validity, this short scale ought to be accompanied with

other personality scales (Balgiu, 2018) whenever it is possible, otherwise interpretation of the results should be made with caution.

Conclusions

This study contributes to increase knowledge about the validity of the BFI-10 in a representative sample of older adults. And it highlights possible method effects associated with the valence of the items to improve the interpretation of the psychometric properties of the scale. It also warns to consider these effects particularly when studying personality in older people.

be accessed at the SHARE Research Data Center to the entire research community free of charge (www.share-project.org). Data are available from the authors upon reasonable request and with permission of the SHARE project.

Conflicts of Interest: The authors declare that there is no conflict of interest.

References

Anglim, J., Horwood, S., Smillie, L. D., & Wood, J. K. (2020). Supplemental Material for Predicting Psychological and Subjective Well-Being From Personality: A Meta-Analysis. *Psychological Bulletin, 146*(4), 279–323. https://doi.org/10.1037/bul0000226.supp

Arias, V. B., Garrido, L. E., Jenaro, C., Martínez-Molina, A., and Arias, B. (2020). A little garbage in, lots of garbage out: assessing the impact of careless responding in personality survey data. *Behavioral Research Methods 52*, 2489–2505. doi: 10.3758/ s13428-020-01401-8

Atherton, O. E., Sutin, A. R., Terracciano, A., & Robins, R. W. (2022). Stability and

    Change in the Big Five Personality Traits: Findings From a Longitudinal Study of

    Mexican-Origin Adults. *Journal of Personality and Social Psychology*, *122*(2),

    337–350. https://doi.org/10.1037/pspp0000385

Balgiu, B. A. (2018). The psychometric properties of the Big Five inventory-10 (BFI-

    10) including correlations with subjective and psychological well-being. *Global*

    Journal of Psychology Research: New Trends and Issues, 8(2), 61–69.

    https://doi.org/10.18844/gjpr.v8i2.3434

Berg, A. I., & Johansson, B. (2014). Personality change in the oldest-old: Is it a matter

    of compromised health and functioning? *Journal of Personality, 82*(1), 25-31.

Bergmann, M., Scherpenzeel, A., & Börsch-Supan, A. (2019). *SHARE Wave 7*

    Methodology: Panel innovations and life histories. Max Plank Institute for

    Social Law and Social Policy, Munich Center for the Economics of Aging.

Börsch-Supan, A. (2020). *Survey of Health, Ageing and Retirement in Europe (SHARE)*

    Wave 7. Release version: 7.1.1. SHARE-ERIC. Data set.

    https://doi.org/10.6103/SHARE.w7.711

Börsch-Supan, A., M. Brandt, C. Hunkler, T. Kneip, J. Korbmacher, F. Malter, B.

    Schaan, S. Stuck, & S. Zuber (2013). Data Resource Profile: The Survey of

    Health, Ageing and Retirement in Europe (SHARE). *International Journal of*

    Epidemiology, *42*(4), 991-1001. https://doi.org/10.1093/ije/dyt088

Brandt, N. D., Becker, M., Tetzner, J., Brunner, M., Kuhl, P., & Maaz, K. (2020).

    Personality Across the Lifespan. *European Journal of Psychological Assessment*,

    36(1), 162–173. https://doi.org/10.1027/1015-5759/a000490

Buecker, S., Maes, M., Denissen, J. J. A., & Luhmann, M. (2020). Loneliness and the Big Five Personality Traits: A Meta-analysis. *European Journal of Personality, 34*(1), 8–28. https://doi.org/10.1002/per.2229

Carciofo, R., Yang, J., Song, N., Du, F., & Zhang, K. (2016). Psychometric Evaluation of Chinese-Language 44-Item and 10-Item Big Five Personality Inventories, Including Correlations with Chronotype, Mindfulness and Mind Wandering. *PLOS ONE, 11*(2), e0149963.

https://doi.org/10.1371/JOURNAL.PONE.0149963

Costa, P. T., & McCrae, R. R. (1992). *NEO-PI-R Professional manual*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., McCrae, R. R., & Löckenhoff, C. E. (2019). Personality Across the Life Span. *Annual Review of Psychology*, *70*, 423–448. https://doi.org/10.1146/annurev-psych-010418-103244

Courtois, R., Petot, J. M., Plaisant, O., Allibe, B., Lignier, B., Réveillère, C., Lecocq, G., & John, O. (2020). Validation of the French version of the 10-item Big Five Inventory. *Encephale, 46*(6), 455–462.

https://doi.org/10.1016/J.ENCEP.2020.02.006

Damian, R. I., Spengler, M., Sutu, A., & Roberts, B. W. (2019). Sixteen going on sixty-six: A longitudinal study of personality stability and change across 50 years. *Journal of Personality and Social Psychology*, *117*(3), 674–695. https://doi.org/10.1037/pspp0000210

Deng, L., & Chan, W. (2017). Testing the Difference Between Reliability Coefficients Alpha and Omega. *Educational and Psychological Measurement, 77*(2), 185–203. https://doi.org/10.1177/0013164416658325

Dong, Y., & Dumas, D. (2020). Are personality measures valid for different populations? A systematic review of measurement invariance across cultures, gender, and age. *Personality and Individual Differences*, *160*(March), 109956. https://doi.org/10.1016/j.paid.2020.109956

Eisinga, R., Grotenhuis, M. Te, & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health, 58*(4), 637–642. https://doi.org/10.1007/s00038-012-0416-3

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

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. Jr., (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

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

Hensley, B., Martin, P., Margrett, J. A., MacDonald, M., Siegler, I. C. & Poon, L. W. (2012). Life events and personality predicting loneliness among centenarians: Findings from the Georgia Centenarian Study. *The Journal of Psychology, 146*, 173–188.

Hofmans, J., Kuppens, P., & Allik, J. (2008). Is short in length short in content? An examination of the domain representation of the Ten Item Personality Inventory scales in Dutch language. *Personality and Individual Differences, 45*(8), 750–755. https://doi.org/10.1016/j.paid.2008.08.004

Hrebícková, M., Jelínek, M., Blatný, M., Brom, C., Buresová, I., Graf, S., Mejzlíková, T., Vazsonyi, A. T., & Zábrodská, K. (2016). Big Five Inventory: Basic psychometric properties of the Czech version of BFI-44 and BFI-10. *Ceskoslovenska Psychologie, 60*(6), 567–583.

Hu, L., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modelling, 6*(1), 1-55. http://dx.doi.org/10.1080/10705519909540118

Hughes, M. E., Waite, L. J., Hawkley, L. C., & Cacioppo, J. T. (2004). A short scale for measuring loneliness in large surveys results from two population-based studies. *Research on Aging, 26*(6), 655-672. http://doi.org/10.1177/0164027504268574

Hyde, M., Wiggins, R. D., Higgs, P., & Blane, D. B. (2003). A measure of quality of life in early old age: The theory, development and properties of a needs satisfaction model (CASP-19). *Aging and Mental Health, 7*(3), 186–194. https://doi.org/10.1080/1360786031000101157

Itzick, M., Kagan, M., & Zychlinski, E. (2020). The Big Five Personality Traits as Predictors of Loneliness among Older Men in Israel. *Journal of Psychology: Interdisciplinary and Applied, 154*(1), 60–74. https://doi.org/10.1080/00223980.2019.1653250

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

John, R. K., Xavier, B., Waldmeier, A., Meyer, A., & Gaab, J. (2019). Psychometric

    evaluation of the BFI-10 and the NEO-FFI-3 in Indian adolescents. *Frontiers in*

    Psychology, 10, 1057. https://doi.org/10.3389/fpsyg.2019.01057

Kail, R., and Y.-s. Park. 1990. Impact of practice on speed of mental rotation. Journal

    ofExperimental Child Psychology 49:227–44.

Kandler, C., Kornadt, A. E., Hagemeyer, B., & Neyer, F. J. (2015). Patterns and sources

    of personality development in old age. *Journal of Personality and Social*

    Psychology, 109(1), 175.

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. The

    Guilford Press.

Lachmann, B., Sariyska, R., Kannen, C., Blaszkiewicz, K., Trendafilov, B., Andone, I.,

    Eibes, M., Markowetz, A., Li, M., Kendrick, K. M., & Montag, C. (2018).

    Contributing to overall life satisfaction: Personality traits versus life satisfaction

    variables revisited-is replication impossible? *Behavioral Sciences, 8*(1), 1–24.

    https://doi.org/10.3390/bs8010001

Lechner, C. M., & Rammstedt, B. (2015). Cognitive ability, acquiescence, and the

    structure of personality in a sample of older adults. *Psychological Assessment*,

    27(4), 1301–1311. https://doi.org/10.1037/pas0000151

Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for

    multivariate measurement and modeling with latent variables: When "good"

    indicators are bad and "bad" indicators are good. *Psychological Methods, 4*(2),

    192–211. https://doi.org/10.1037/1082-989X.4.2.192

Marsh, H. W., Hau, K. T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research, 33*(2), 181–220. https://doi.org/10.1207/s15327906mbr3302_1

McCrae, R. R., & Costa, P. T. (2004). A contemplated revision of the NEO Five-Factor Inventory. *Personality and Individual Differences, 36*(3), 587–596. https://doi.org/10.1016/S0191-8869(03)00118-1

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates Publishers.

Menold, N. (2020). How Do Reverse-keyed Items in Inventories Affect Measurement Quality and Information Processing? Field Methods, 32(2), 140–158. https://doi.org/10.1177/1525822X19890827

Mõttus, R., Johnson, W., & Deary, I. J. (2012). Personality traits in old age: measurement and rank-order stability and some mean-level change. *Psychology and aging, 27*(1), 243.

Muthén, L. K. & Muthén, B. O. (1998-2017). *Mplus User's Guide, 8th ed.* Muthén & Muthén.

Olaru, G., & Allemand, M. (2021). Correlated personality change across time and age. *European Journal of Personality*, *0*(0), 089020702110140. https://doi.org/10.1177/08902070211014054

Oliver, A., Sentandreu-Mañó, T., Tomás, J. M., Fernández, I., & Sancho, P. (2021). Quality of life in european older adults of share wave 7: Comparing the old and

the oldest-old. *Journal of Clinical Medicine, 10*(13), 2850.

https://doi.org/10.3390/jcm10132850

Rammstedt, B., & Farmer, R. F. (2013). The impact of acquiescence on the evaluation

of personality structure. *Psychological Assessment*, *25*, 1137– 1145.

http://dx.doi.org/10.1037/a0033323

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A

10-item short version of the Big Five Inventory in English and German. *Journal*

of Research in Personality, 41(1), 203–212.

https://doi.org/10.1016/j.jrp.2006.02.001

Rammstedt, B., Danner, D., Soto, C. J., & John, O. P. (2020). Validation of the Short

and Extra-Short Forms of the Big Five Inventory-2 (BFI-2) and Their German

Adaptations. *European Journal of Psychological Assessment, 36*(1), 149–161.

https://doi.org/10.1027/1015-5759/a000481

Rammstedt, B., Kemper, C. J., Klein, M. C., Beierlein, C., & Kovaleva, A. (2013). A

Short Scale for Assessing the Big Five Dimensions of Personality: 10 Item Big

Five Inventory (BFI-10). *Methods, Data, Analyses, 7*(2), 233-249.

https://doi.org/10.12758/MDA.2013.013

Russel, D., Peplau, L. A. & Cutrona, C. E. (1980). The revised UCLA Loneliness Scale:

concurrent and discriminant validity evidence. *Journal of Personality and Social*

Psychology, 39(3), 472-480. https://doi.org/10.1037/0022-3514.39.3.472

Salazar, M. S. (2015). El dilema de combinar ítems positivos y negativos en escalas.

Psicothema, 27(2), 192–199. https://doi.org/10.7334/psicothema2014.266

Saucier, G. (1994). Mini markers: A brief version of the Goldberg's unipolar Big Five

markers. *Journal of Personality Assessment, 63*(3), 506–516.

https://doi.org/10.1207/s15327752jpa6303_8

Schutter, N., Koorevaar, L., Holwerda, T. J., Stek, M. L., Dekker, J., & Comijs, H. C.

(2020). "Big Five" personality characteristics are associated with loneliness but

not with social network size in older adults, irrespective of depression.

International Psychogeriatrics, 32(1), 53–63.

https://doi.org/10.1017/S1041610219000231

Seng Kam, C. C., & Meyer, J. P. (2015). Implications of Item Keying and Item Valence

for the Investigation of Construct Dimensionality. *Multivariate Behavioral

Research, 50*(4), 457–469. https://doi.org/10.1080/00273171.2015.1022640

Soto, C. J. & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and

assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and

predictive power. *Journal of Personality and Social Psychology, 113,*117-143.

doi: 10.1037/pspp0000096.

Steinmann, I., Strietholt, R., and Braeken, J. (2021). A constrained factor mixture

analysis model for consistent and inconsistent respondents to mixed-worded

scales. *Psychological Methods 1*–36. doi: 10.1037/met0000392

Suárez-Alvarez, J., Pedrosa, I., Lozano, L. M., García-Cueto, E., Cuesta, M., & Muñiz,

J. (2018). Using reversed items in likert scales: A questionable practice.

Psicothema, 30(2), 149–158. https://doi.org/10.7334/psicothema2018.33

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K.

A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 10-39).

Sage.

Terracciano, A., Costa, P. T., Jr., & McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin, 32*(8), 999–1009. https://doi.org/10. 1177/0146167206288599

Thalmayer, A. G., Saucier, G., & Eigenhuis, A. (2011). Comparative Validity of Brief to Medium-Length Big Five and Big Six Personality Questionnaires. *Psychological Assessment, 23*(4), 995–1009. https://doi.org/10.1037/a0024165

van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of Reverse Wording of Questionnaire Items: Let's Learn from Cows in the Rain. *PLoS ONE, 8*(7), 1–7. https://doi.org/10.1371/journal.pone.0068967

Wagner, J., Lu¨ dtke, O., & Robitzsch, A. (2019). Does personality become more stable with age? Disentangling state and trait effects for the big five across the life span using local structural equation modeling. *Journal of Personality and Social Psychology. 116*(4), 666–680. https://doi.org/10.1037/pspp0000203

Wagner, J., Ram, N., Smith, J., & Gerstorf, D. (2016). Personality trait development at the end of life: Antecedents and correlates of mean-level trajectories. *Journal of Personality and Social Psychology, 111*(3), 411.

Wang, B., & Dong, X. (2018). The Association Between Personality and Loneliness: Findings From a Community-Dwelling Chinese Aging Population. *Gerontology and Geriatric Medicine, 4*, 1-9. https://doi.org/10.1177/2333721418778181

Woods, S. A., & Hampson, S. E. (2005). Measuring the big five with single items using a bipolar response scale. *European Journal of Personality, 19*, 373–390. https://doi.org/10.1002/per.542

Table 1. Fit indices of the tested models.

Model	χ^2	df	p	RMSEA	90%CI	CFI	SRMR
M1. Five-factor model	9129.60	25	< .001	.070	.068 - .071	.90	.032
M2. Five-factor model & M.E. factor for reversed items	7021.82	20	< .001	.068	.067 - .070	.93	.028
M3. Five-factor model & M.E. factor for positive items	5691.83	20	< .001	.061	.060 - .063	.94	.026
M4. Five-factor model & both M.E. factors	3225.13	15	< .001	.053	.052 - .055	.97	.019

M = Model; M.E. = Method effect

Table 2. Correlation matrix between the personality dimensions considering methods effects into the model and without method effects. Note: Estimates when method effects are considered in the left, and without method effects in the right; all correlations statistically significant (p< .05)

	O	C	E	A	N
Openness (O)	1				
Conscientiousness (C)	.20/.34	1			
Extraversion (E)	.22/.39	.44/.49	1		
Agreeableness (A)	-.20/.15	.38/.57	.58/.72	1	
Neuroticism (N)	-.35/-.18	-.44/-.33	-.55/-.42	-.93/-.68	1

Table 3. Correlations between the personality dimensions and loneliness, well-being and life satisfaction. All correlations statistically significant ($p < .05$).

	Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness
Loneliness	-.174	-.079	-.154	.267	-.091
Well-being	.243	.159	.234	-.366	.166
Life satisfaction	.163	.119	.136	-.254	.099

Figure 1. Best-fitting model structure and standardized factor loadings. All relationships statistically significant ($p < .05$) unless stated otherwise. Note: O = Openness; C = Conscientiousness; E = Extraversion; A = Agreeableness; N = Neuroticism; V. E. = Valence effect; ns = not significant.