

Brandt, Naemi; Becker, Michael; Tetzner, Julia; Brunner, Martin; Kuhl, Poldi; Maaz, Kai
Personality across the lifespan. Exploring measurement invariance of a short big five inventory from ages 11 to 84

formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:

formally and content revised edition of the original source in:

European journal of psychological assessment 36 (2020) 1, S. 162-173



Bitte verwenden Sie in der Quellenangabe folgende URN oder DOI /
Please use the following URN or DOI for reference:

urn:nbn:de:0111-pedocs-237822

10.25656/01:23782

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-237822>

<https://doi.org/10.25656/01:23782>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz:
<http://creativecommons.org/licenses/by-nc/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen, solange Sie den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen und das Werk bzw. den Inhalt nicht für kommerzielle Zwecke verwenden.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License:
<http://creativecommons.org/licenses/by-nc/4.0/deed.en> - You may copy, distribute and render this document accessible, make adaptations of this work or its contents accessible to the public as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work, provided that the work or its contents are not used for commercial purposes.

By using this particular document, you accept the above-stated conditions of use.



Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

Accepted manuscript version (after peer review) of the following article:

Brandt, N.D., Becker, M., Tetzner, J., Brunner, M., Kuhl, P., & Maaz, K. (2020). Personality across the lifespan: Exploring measurement invariance of a Short Big Five Inventory from ages 11 to 84. *European Journal of Psychological Assessment*, 36(1), 162–173.
<https://doi.org/10.1027/1015-5759/a000490>

© 2018 Hogrefe Publishing

This version of the article may not completely replicate the final version published in the journal. It is not the version of record and is therefore not suitable for citation.

The accepted manuscript is subject to the Creative Commons licence CC-BY-NC.

Personality across the Lifespan: Exploring Measurement Invariance of a Short Big Five Inventory from Ages 11 to 84

Naemi D. Brandt^{1,2}, Michael Becker^{1,2}, Julia Tetzner^{1,2}, Martin Brunner³, Poldi Kuhl⁴, Kai Maaz¹

¹German Institute for International Educational Research, Berlin/Frankfurt a.M.

²Leibniz Institute for Science and Mathematics Education, Kiel

³University of Potsdam, Potsdam

⁴Leuphana University, Lüneburg

Correspondence address: Naemi D. Brandt, German Institute for International Educational Research, Warschauer Straße 34-38, 10243 Berlin. brandt@dipf.de.

Author note

We are grateful to the editor and an anonymous reviewer for valuable feedback. We thank Johannes Hartig for helpful comments on this project. Furthermore, we thank Holly Painter for editorial assistance.

Article type: Multistudy report

Word count: 7477

Tables:

Table 1: Measurement Invariance of Multi-Group CFA Models from Ages 11 to 84

Table 2: Item Labels and Intercepts (Model 3f Table 1).

Table 3: Loadings of BFI-S Personality Items to Big Five Factors: Results of the Partial Scalar Invariance Model (Model 3f Table 1)

Figures:

Figure 1: Schematic illustration of an exploratory structural equation model (ESEM).

Figure 2: Correlation patterns of Big Five latent factors in all age groups. (Results of Model 3f Table 1)

Abstract

Personality is a relevant predictor for important life outcomes across the entire lifespan. Although previous studies have suggested the comparability of the measurement of the Big Five personality traits across adulthood, the generalizability to childhood is largely unknown. The present study investigated the structure of the Big Five personality traits assessed with the Big Five Inventory-SOEP Version (BFI-S) across a broad age range spanning 11 to 84 years. We used two samples of $N = 1090$ children (52 % female, mean age = 11.87) and $N = 18789$ adults (53 % female, mean age = 51.09), estimating a multi-group CFA analysis across four age groups (late childhood: 11 to 14 years; early adulthood: 17 to 30 years; middle adulthood: 31 to 60 years; late adulthood: 61 to 84 years). Our results indicated the comparability of the personality trait metric in terms of general factor structure, loading patterns, and the majority of intercepts across all age groups. Therefore, the findings suggest both a reliable assessment of the Big Five personality traits with the BFI-S even in late childhood and a vastly comparable metric across age groups.

Keywords: personality traits, measurement invariance, ESEM, lifespan, late childhood

Personality across the Lifespan: Exploring Measurement Invariance of a Short Big Five Inventory from Ages 11 to 84

Previous research has frequently shown that personality traits have a substantial influence on different life domains. They are meaningful for academic success, health, and well-being, amongst other domains (Anglim & Grant, 2016; Poropat, 2009; Sirois & Hirsch, 2015). Moreover, one important finding of recent research is that personality traits do not remain entirely stable throughout life and that they are related to the experience of different life events (Lüdtke, Roberts, Trautwein, & Nagy, 2011; Roberts & DelVecchio, 2000; Specht, Egloff, & Schmukle, 2011). Studying those dynamics over the life course brings new challenges including the measurement of personality traits (Milfont & Fischer, 2010).

In order to make assumptions about changes in personality and their impact on relevant life outcomes, it is necessary to investigate whether personality traits can be assessed validly in a similar way across different age groups. So far, most previous studies have either examined the structure of personality for isolated age groups or small age ranges longitudinally (e.g., Asendorpf & van Aken, 2003; John, Caspi, Robins, Moffitt, & Stouthamer-Loeber, 1994; Measelle, John, Ablow, Cowan, & Cowan, 2005) or excluded childhood from cross-sectional multi-group analyses across the lifespan (e.g., Marsh, Nagengast, & Morin, 2013). The use of different personality inventories for different age groups also limits comparisons.

Therefore, the main aim of the present study was to investigate the psychometric properties of a short personality inventory across the lifespan, specifically in terms of measurement invariance across different age groups. We focused on a short personality inventory, which has many advantages, such as test efficiency in large-scale surveys and panel studies. Personality traits assessed with short instruments based on the Big Five Inventory (BFI; John, Donahue, & Kentle, 1991; John & Srivastava, 1999) are useful for many disciplines (such as psychology, educational science, and economics) in order to

explain individual differences, for example, in educational outcomes and returns (e.g., Caliendo, Fossen, & Kritikos, 2014; Marsh et al., 2013; Specht et al., 2011). So far, researchers have primarily tested the measurement properties of short BFI inventories in adult samples. Therefore, the present study was – to the best of our knowledge – the first to investigate the psychometric properties of a short BFI Inventory from late childhood to late adulthood.

Personality Structure across the Lifespan

Personality traits are individual characteristics of a person that have an impact on his or her experiences and behavior (McCrae & Costa, 2008). Initially, researchers assumed that personality traits were stable and fully developed by the age of 30 with few changes after that (Costa & McCrae, 1997). However, current research suggests that personality develops across the entire lifespan (Roberts, Walton, & Viechtbauer, 2006). On the one hand, personality development follows a normative trajectory, and consistency rises with increasing age, also known as the cumulative continuity principle (Caspi, Roberts, & Shiner, 2005). On the other hand, life events and conditions (e.g., educational transitions, illness, or unemployment) are substantively related to the course of personality development and changes in personality (Lüdtke et al., 2011; Specht et al., 2011).

For adulthood, the most commonly used model to describe personality is the Five-Factor model (Big Five), which uses five broad factors to describe individual differences in experience and behavior (John, Naumann, & Soto, 2008): openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. These five factors build the highest order of a hierarchical personality model that subsumes narrower facets covering the diversity of human beings (John et al., 2008). Personality theory supposes the Big Five to be nearly uncorrelated, while on the facet level, correlations may occur (Costa, & McCrae, 1995). Empirically, researchers have often found significant relations between factors (Ashton, Lee, Goldberg, & De Vries, 2009). Some authors interpret this as evidence of

higher-order factors above the Big Five (DeYoung, 2006; van der Linden, te Nijenhuis, & Bakker, 2010). Others attribute identified relations to artificial response tendencies (Biesanz & West, 2004; Chang, Connelly, & Geeza, 2012).

While researchers generally use the Big Five to describe personality in adulthood, it is not yet well known whether this model is also applicable to children's personality structures. This knowledge gap might be due to a research tradition developed largely independently of the Big Five approach. In this research tradition, researchers describe individual differences in children not as personality but as temperament (Zentner & Bates, 2008). This line of research is exclusively focused on childhood and therefore offers few links between adults' and children's personality models (De Fruyt et al., 2000). However, more recent studies have shown theoretical and empirical overlap between temperament and the Big Five (Caspi & Shiner, 2006; De Pauw, Mervielde, & Van Leeuwen, 2009; Rothbart, Ahadi, & Evans, 2000). This might indicate that the Big Five personality model is applicable to childhood. Researchers developed Big Five instruments with easy language for children (e.g., HiPIC; Bleidorn & Ostendorf, 2009), which confirmed a five-factor structure as established in adulthood. Theoretically unintended loadings also occurred. There are also further differences in the operationalization of child-specific inventories compared to adult inventories. Openness, for example, is conceptualized as imagination, omitting aspects of aesthetic and artistic interests in childhood inventories (Bleidorn & Ostendorf, 2009). Moreover, the Big Five showed stronger interrelations in childhood than in adulthood. It is still unclear whether these differences in metric and structure between age groups occur due to developmental processes, or if the differences are evoked by the use of inventories with slightly different conceptualizations or wordings (De Pauw et al., 2009).

In order to separate methodological issues from developmental processes, some studies have used adult Big Five Inventories (NEO-FFI; Costa & McCrae, 1992) with children and compared the structure to other age groups (Allik, Laidra, Realo, & Pullmann,

2004) or investigated age differences in the personality structure longitudinally (McCrae et al., 2002). These studies showed that the Big Five were apparent, and to some extent comparable to the adult Big Five structure, by age 12. Even though differences in factor structures of adults and children remained, Allik et al. (2004) found no evidence for additional factors besides the Big Five using the NEO-FFI. There were similar results when researchers investigated the structure using the Big Five Inventory: In a sample of 10- to 20-year olds, Soto, John, Gosling, and Potter (2008) found five factors while controlling for acquiescent response style. With increasing age, more items loaded on the theoretically intended domain. When the researchers controlled for acquiescence, the five-factor structure was already very recognizable at age 10. The explained variance of factors, though, was smallest in childhood and increased during adulthood. Thus, the results obtained from using adult inventories for samples of children provide evidence for the reproducibility of the Big Five personality structure in childhood with longer Big Five inventories. However, previous studies could not contrast children's personality structures with adults' psychometric properties across the entire lifespan and only contrasted them with the specific and frequently studied young adult age group.

Panel survey designs are suitable for investigating personality traits and their relations with relevant life outcomes across a broad age range or even across the entire lifespan. The large samples assessed in these contexts necessitate efficient inventories, so researchers often use short personality inventories. The generalizability from longer tests to shorter ones must be investigated when including new research questions (Ziegler, Poropat, & Mell, 2014). A first study by Marsh et al. (2013) examined the psychometric properties of a short Big Five inventory (BFI) with 15 items in a British survey sample of 15- to 99-year olds. The authors found a comparable metric across age groups: The relations of the Big Five, the item loadings, and the vast majority of item intercepts were similar when the authors allowed for correlations between negatively worded items. In Germany, using data from the German Socioeconomic

Panel (SOEP), researchers confirmed longitudinal measurement invariance (Specht et al., 2011) and also multi-group invariance (Lucas & Donnellan, 2011) of the SOEP version of the Big Five Inventory (BFI-S) across adulthood when they modeled personality domains separately. So far, no study has investigated whether this pattern of results remains when surveying children using a German short BFI.

The Present Study

To date there is little systematic research on the comparability of personality traits from late childhood to late adulthood. Although there is evidence for longer Big Five inventories (e.g., Allik et al., 2004; McCrae et al., 2002; Soto et al., 2008), little is known about short inventories. Therefore, the aim of the present study was to investigate psychometric properties of a short Big Five inventory by means of three aims.

First, we considered the personality structure in childhood and adulthood (Aim 1). Since previous studies based on childhood and adult Big Five inventories supported the existence of a five-factor structure in childhood, we expected to find the general structure of five factors across the entire considered age span, using the short Big Five inventory.

Besides establishing the Big Five structure, the second aim was to investigate the comparability of the personality trait metric across different age groups (Aim 2). Previous research has found early indications of similarity for longer inventories in childhood (Allik et al., 2004; McCrae et al., 2002; Soto et al., 2008), pointing to a loading pattern that is somewhat more ambiguous. For short inventories, only one study has investigated comparability from adolescence to old age (Marsh et al., 2013); it indicated comparable loading patterns as well as widely comparable intercepts. We therefore analyzed how far the personality metric is comparable when considering childhood, too.

Finally, in a third step, we addressed the comparability of the Big Five interrelations (Aim 3). Previous studies have demonstrated substantial correlations between personality factors in childhood (Allik et al., 2004; McCrae et al., 2002) as well as in adulthood (Ashton

et al., 2009). Furthermore, relations between the Big Five factors varied by age from adolescence to late adulthood (Marsh et al., 2013). We likewise expected to find variability in correlational patterns from childhood to late adulthood.

Method

Sample and Participants

We used two samples, one of children and one of adults, covering an age range from 11 to 84 years. The sample of children was based on data from the KEGS project (development of competencies in primary school, Fuchs & Brunner, 2014) and included $N = 1090$ sixth graders (52 % female, age: $M = 11.87$, $SD = .56$, $Mdn = 12$, Range = 11-14 years) from 68 randomly drawn primary schools in the German federal state of Brandenburg in 2011. Trained test administrators administered the survey. The students filled in the questionnaires in their classrooms on their own.

The sample of adults was from the Socio-Economic Panel in Germany (Socio-Economic Panel (SOEP) 2016; Wagner, Frick, & Schupp, 2007). A representative survey of households in Germany, the SOEP includes questions regarding the economic situation of household members, as well as questions about psychosocial life conditions. Households were chosen using a multistage randomized sampling strategy. The sample included $N = 18789$ adults (53 % female, age: $M = 51.09$, $SD = 17.42$; $Mdn = 52$, Range: 17-84 years). Data was collected in 2013 primarily via online or paper-pencil surveys. Trained interviewers personally surveyed about 15% of the adult respondents.

Instrument

In both samples, we used the Big Five Inventory SOEP-version (BFI-S; Gerlitz & Schupp, 2005; Lang, 2005) to assess personality. The BFI-S is a self-report inventory, originally based on a German translation of the BFI-44 by John et al. (1991). This short version was developed for the survey design of the SOEP. For information about reliabilities

based on evaluations of the BFI-S within SOEP and convergent validities, see Gerlitz and Schupp (2005), Hahn, Gottschling, and Spinath (2012) as well as Lang (2005).

The BFI-S assesses four of the personality traits – conscientiousness, extraversion, agreeableness, and neuroticism – with three items each (including one reverse-coded item for each). Due to heterogeneity of openness to experience, there are four items for this fifth trait (Lang, 2005), none of which are reverse-coded. For original item wording, see Table 2. Respondents rate all 16 items on a 7-point Likert scale from 1 (*doesn't apply at all*) to 7 (*applies perfectly*). We estimated reliabilities of scale values for the full sample using the model-based reliability index ω (McDonald, 1999). Analyses showed low to satisfactory values for children and adults, respectively: ω (conscientiousness) = .69/.63, ω (agreeableness) = .75/.69, ω (extraversion) = .49/.68, ω (openness to experience) = .76/.68, and ω (neuroticism) = .57/.67.

Statistical Approach

Previous studies most frequently used exploratory factor analyses or principal component analyses to examine the Big Five factor structure, while confirmatory factor analyses (CFA) often failed to establish the assumed factor structure (Church & Burke, 1994; Vassend & Skrondal, 1997). Researchers often attributed this to the strong assumption of simple structure within the independence cluster model (ICM) of a confirmatory approach (Marsh et al., 2010). Researchers often modeled Big Five domains separately to avoid problems that arose while investigating the Big Five factor structure by applying an ICM (e.g., Lucas & Donnellan, 2011; Specht et al., 2011). However, Asparouhov and Muthén (2009) proposed a combined approach based on exploratory rotation principles and on structural equation modeling to investigate comparability of psychometric properties across the lifespan (exploratory structural equation modeling, ESEM). Within the measurement model, the exploratory part overcomes restrictions of zero cross-loadings, while flexibility of

structural equation modeling allows researchers to test models directly (e.g., measurement invariance in multi-group models).

In our study, we used a cross-validation strategy combining ESEM (see Figure 1 for a schematic ESEM model) with multi-group mean and covariance structure (MGMCS) measurement invariance testing within the CFA framework. We therefore split our sample into two halves and created a multi-group ESEM model to identify cross-loadings in one half-sample ($n = 9968$). We then included all statistically significant ($p < .01$) non-zero cross-loadings in a CFA model¹ and investigated measurement invariance using the other half sample ($n = 9911$). We constructed four age groups: late childhood (11-14 years, $n = 547$), early adulthood (17-30 years, $n = 1507$), middle adulthood (31-60 years, $n = 4753$), and late adulthood (61-84 years, $n = 3104$). For comparability to the study of Marsh et al. (2013), we used a similar age categorization system. As a check of robustness and in order to deal with problems in breaking down a continuous variable such as age into discrete clusters, we re-estimated our models using varying age-clusters and additionally applied a continuous modeling approach (local structural equation modeling, LSEM; Hildebrandt, Lüdtke, Robitzsch, Sommer, & Wilhelm, 2016; Hildebrandt, Wilhelm, & Robitzsch, 2009) to describe courses of loadings and intercepts across the age range. Results from all additional analyses are in the online supplemental material.

To examine the factor structure of personality across the considered age range (Aim 1), we evaluated a configural CFA model using well-established model fit criteria (CFI $> .95$ -.97; RMSEA $< .05$ -.08, SRMR $< .05$ -.10, and AIC; Hu & Bentler, 1999; Schermelleh-Engel, Moosbrugger, & Müller, 2003).

To test measurement invariance of personality assessment from late childhood to late adulthood (Aims 2 and 3), we specified increasingly restrictive multi-group CFA models. For this purpose, we tested metric invariance (equal loadings across groups), scalar invariance

¹ Except for the cross-loading of item A3 to conscientiousness because this resulted in a negative residual variance of that item.

(adding equal intercepts across groups), and structural invariance (adding equal variances and covariances of latent factors across groups) against each other across all age groups (Widaman & Reise, 1997). As we cannot assume that unsystematic error influences are the same across age-groups, we did not test for strict invariance (equal residual variances; Little, 2013).

We also allowed residual variances of negatively worded items to correlate between personality domains in all models. Although *a priori* correlated error variances should only be specified if there is a substantive rationale for doing so (Marsh et al., 2013), we assumed that this was the case in our study. Previous studies have often found systematic response tendencies to reverse-coded or negatively worded items for children (Marsh, 1986) and also for adults (Rammstedt & Farmer, 2013). Therefore, the similarity of responses to negatively worded items of different factors may simply result from their modified phrasing. We identified scales of latent variables by fixing the variance of the first group to one and the mean to zero.

We evaluated the specified models using changes in model fit criteria like CFI, RMSEA, and SRMR. With regard to the invariance testing of loadings, according to Chen (2007) and Cheung and Rensvold (2002), a non-significant model deterioration is indicated by a decrease of less than .010 in CFI, an increase of less than .015 in RMSEA, or an increase of less than .030 in SRMR. With respect to the invariance of item intercepts, in addition to the same cut-offs for CFI and RMSEA, the SRMR should not increase by more than .010. We tested partial invariance of parameters when the model fit deterioration was significant according to these rules. We then checked modification indices (Lagrange multipliers), freed the equality constraint with the highest value, and fit the model again. We continued doing so until model fit deterioration remained within acceptable ranges according to Chen (2007). We also report scaled χ^2 difference tests for the sake of completeness but did not use them for

model evaluation mainly because of their well-known sensitivity to trivial differences between specified models and empirical data.

We used the full information maximum likelihood (FIML) approach with a robust maximum likelihood estimator (MLR) to account for both missing data and significant skew and kurtosis of item responses in all age groups. To account for possible biases resulting from nested data structure (students nested in schools, adults nested in households), we performed additional analyses. The pattern of results of these analyses was comparable to results derived without considering clustering of data and is therefore not reported in detail. We used *Mplus* 7.4 (Muthén & Muthén, 1998-2015) for ESEM modeling and R (R Core Team, 2016) for CFA models, LSEM, and descriptive results.

Results

The first aim of our study was to investigate whether personality traits from late childhood to late adulthood can be described with the Big Five model. As Table S1 (online supplemental material) shows, we found first descriptive indications of the correlational pattern of personality items. In both samples (KEGS and SOEP), there were significant relations between items of a theoretically intended personality domain, but we also found significant, albeit mostly smaller, loadings of items to theoretically unintended personality domains. (Tables S2 and S3 in the online supplemental material show correlations of personality items in adult subgroups).

As a basis for cross-validation within the CFA framework, we first estimated a configural ESEM model, which fits the data well with $\chi^2(167) = 601.86, p < .001$; AIC = 529454.23, CFI = .982, RMSEA = .032, 90% CI = [.030, .035], SRMR = .014 (see Table S4 in the online supplemental material for loadings of the ESEM model). To gain more insight into the results in the context of our first aim, we evaluated the model fit of the configural CFA model with five factors based on the ESEM model (see measurement section). The configural CFA model fit the data well across the whole age span (see Table 1). We improved

the model fit when we allowed residual variances of negatively worded items to correlate (model without residual correlations: $\Delta\chi^2 (\Delta 36) = 543.89$, $p < .001$; CFI = .960, RMSEA = .047, 90% CI = [.044, .049], SRMR = .023). Finally, we checked whether residual correlations between negatively worded items were comparable across the considered age range by fixing them to equality. Model fit suggested comparability ($\chi^2 (228) = 666.17$, $p < .001$, CFI = .981, RMSEA = .033, 90% CI = [.030, .036], SRMR = .017, range of residual correlations: -.31 - .21).

Furthermore, our second aim was to test the comparability of the personality trait metric across the age span. Therefore, we tested metric and scalar invariance. Results demonstrated comparable measurement properties of the BFI-S across all age groups. In particular, we confirmed partial scalar invariance across age groups (Table 1): The pattern of loadings, their amount, and the majority (11 out of 16) of item intercepts were comparable across all age groups. Considering modification indices, we found five intercepts in total (C2r, O2, O4, A2, A3), which varied between groups (Table 2). Intercepts of items C2r, O2, and A3 were not invariant in early adulthood (17-30 years), A3 was also not invariant in middle adulthood (31-60 years) and C2r was also not invariant in late adulthood (61-84 years). The intercepts of items O4 and A2 only needed to be freed from equality constraints in childhood (11-14 years). The factor loadings of the resulting model (Table 3) confirmed the theoretically intended pattern: We found the highest loadings of items to their intended domains. The only exception was one significant cross-loading of the reverse-coded item of extraversion (“is reserved”) to agreeableness ($\lambda = -.619$). The robustness checks (see Tables S5 and S6 in the online supplemental material) supported our findings. Results stayed the same across different age categorizations demonstrating partial scalar invariance (except for agreeableness, which showed only one invariant item across age groups). Figures S1-S3 in the supplemental material show patterns of unstandardized loadings, item intercepts, and fit indices as a function of age.

Moreover, our third aim was to test the correlational pattern of personality factors for equality (structural measurement invariance). Descriptive results suggested somewhat comparable relations of Big Five Factors across age groups, but we also discovered differences (see also Figure 2). Although we found significant correlations between Big Five personality traits in all age groups, they were more pronounced in childhood. We also found mixed results regarding the model fit of the structural invariance model (Model 4 in Table 1). The reduction of model fit due to invariant factor correlations was acceptable with respect to most model fit criteria ($\Delta CFI = -.009$, $\Delta RMSEA = .002$). The SRMR, however, increased more strongly (.012) than is considered acceptable by Chen (2007).

Discussion

In this article, our central aim was to investigate the psychometric properties of the BFI-S from late childhood to late adulthood. We built on existing knowledge of psychometric properties indicated by short personality inventories across the lifespan and broadened it by adding children. One first aim was to investigate whether the adult Big Five factor structure is also observable in late childhood. Results from a short BFI inventory suggest that the overall structure of five factors is observable from late childhood to late adulthood. These results are in line with evidence from longer personality inventories that shows the Big Five are already in place in childhood (Allik et al., 2004; McCrae et al., 2002; Soto et al., 2008). Using ESEM, we found item loadings to their theoretically intended personality factor but also additional cross-loadings to unintended domains. Thus, the finding of an incomplete simple structure is in keeping with the difficulties of representing the Big Five factor structure with confirmatory factor analyses (Church & Burke, 1994; Vassend & Skrandal, 1997). However, the ESEM analyses demonstrated that these results are not specific to childhood and similarly show across the entire considered age range. These results could be cross-validated within the CFA framework. Due to its brevity, the BFI-S is not intended to cover a facet structure of personality. In line with theoretical assumptions of personality, simple structure is first and

foremost expected between facets and latent factors (Costa, & McCrae, 1995). We therefore recommend using an ESEM-based approach with cross-validation in the CFA framework when assessing personality with short inventories and predicting external criteria or identifying cross-loadings between domains. Especially when evaluating the joint Big Five model, an ESEM-based procedure within the CFA context could take cross-loadings into account and control for them between groups.

Regarding the second aim of this study, we tested the equivalence of the BFI-S metric for children and adults. We found that the psychometric properties of the measurement model were comparable between groups for the loadings and the vast majority of item intercepts. The group of early adulthood (17-30 years) was responsible for the most non-invariant intercepts (with three non-invariant intercepts), followed by the group of late childhood with two non-invariant intercepts whereas for middle and late adulthood, only one intercept differed statistically from the other age groups. From the constructs' perspective, regarding extraversion and neuroticism full scalar invariance could be established. Concerning openness and conscientiousness, equality constraints had to be relaxed in two groups for one item each. In terms of agreeableness, equality constraints had to be relaxed in three groups for one item each. Therefore, partial scalar invariance could be established for openness, conscientiousness and agreeableness. This implies that the BFI-S allows a comparison of construct-means from late childhood to late adulthood with partial scalar invariance.

Non-invariance of item intercepts could reflect that some behavior or experiences are less frequent in different age groups or have a different meaning in these groups, leading to different item responses. The valence of personality traits may differ along individuals' developmental pathways, as demonstrated by the varying valences of self-concepts, goals, and priorities in life, as well as the achievement of developmental tasks and the need for adaptability (Baltes, Lindenberger, & Staudinger, 2006). For example, the importance of being hardworking or diligent may change across the lifespan. Particularly, being retired may

change the relevance of conscientiousness in older age groups especially from concepts younger adults have at the beginning of their professional careers. In line with this assumption, Specht and colleagues (2011) reported evidence that conscientiousness wanes after people retire. Hence, the intercept of the item “tends to be lazy” differed for older and younger adults. On more methodological grounds, non-invariant items of a construct imply the measurement of fewer items on the same scale and therefore a decrease in the reliability of estimated means because personality domains then rely on fewer comparable items (Steenkamp & Baumgartner, 1998). However, recent studies have suggested that, in particular, non-invariance in both metric and scalar parts of the model may significantly bias results (Guenole & Brown, 2014). Nevertheless, future research needs to assess more thoroughly the conditions under which partial scalar invariance has meaningful consequences for mean or variance comparisons (Putnick & Bornstein, 2016). The result of an age-varying intercept is again not specific to childhood and has also appeared in joint investigations of adolescents and adults (e.g., Marsh et al., 2013). To conclude, while using the BFI-S, researchers could investigate relations of personality with other variables and compare between groups. Moreover, this also allows for an examination of personality development across the life course. Our study therefore adds to existing evidence by demonstrating the comparability of psychometric properties across ages from late childhood to late adulthood.

To address our third aim, we studied the correlational pattern of the Big Five in different age groups. The global model fit of the structural invariance model was mostly satisfactory. For direct model comparisons, we found mixed results. The SRMR clearly increased, which is in the first instance an indicator for wrongly specified latent factor correlations (Hu & Bentler, 1998). We also found differences in factor relations between groups in the descriptive results. In particular, factor relations were smallest for early adulthood. Previous work has referred to less distinct personality factors in childhood (Allik et al., 2004; Soto et al., 2008). The study from Marsh et al. (2013) likewise failed to establish

structural invariance from adolescence to late adulthood. With regard to how researchers might use BFI-S-assessed personality traits to predict and explain external criteria, less distinct factors are not problematic or less predictive *per se* (Booth & Hughes, 2014). However, depending on the research question of interest, researchers should consider interrelations of Big Five factors.

Our study has several strengths: We considered a broad age range, and age groups were immediately comparable in one joint model. Furthermore, we used various modeling procedures to increase confidence in our results. On the other hand, some aspects also limit our approach: Our sample provided responses to the BFI-S using different methods (interview or questionnaire) and in different contexts (in classrooms for KEGS and individually or on a computer for SOEP). Although previous work has shown BFI-S to be invariant across different assessment methods including individual, assisted, or computer-based assessment in adulthood (Lang, John, Lüdtke, Schupp, & Wagner, 2011), we do not know whether the diverse set of methods impacted our results. Therefore, our results could be seen as lower estimates of comparability, and equivalence might be improved if assessment methods did not vary.

Conclusively, our study revealed evidence for the comparability of the personality trait metric of the Big Five from late childhood to late adulthood. We demonstrated this with the BFI-S, a short Big Five inventory originally developed for adults. Therefore, the BFI-S is an inventory that encourages analyzing research questions regarding personality development, its antecedents, correlations, and consequences across a very broad age range or even lifespan.

References

- Allik, J., Laidra, K., Realo, A., & Pullmann, H. (2004). Personality development from 12 to 18 years of age: Changes in mean levels and structure of traits. *European Journal of Personality, 18*(6), 445–462. <https://doi.org/10.1002/per.524>
- Anglim, J., & Grant, S. (2016). Predicting psychological and subjective well-being from personality: Incremental prediction from 30 facets over the Big 5. *Journal of Happiness Studies, 17*(1), 59–80. <https://doi.org/10.1007/s10902-014-9583-7>
- Asendorpf, J.B., & van Aken, M.A.G. (2003). Validity of Big Five personality judgments in childhood: A 9 year longitudinal study. *European Journal of Personality, 17*(1), 1–17. <https://doi.org/10.1002/per.460>
- Ashton, M.C., Lee, K., Goldberg, L.R., & DeVries, 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>
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16*(3), 397–438. <https://doi.org/10.1080/10705510903008204>
- Baltes, P.B., Lindenberger, U., & Staudinger, U.M. (2006). Life span theory in developmental psychology. In W. Damon & R.M. Lerner (Eds.), *Handbook of Child Psychology* (pp. 569–664). New York: Wiley.
- Biesanz, J.C., & West, S.G. (2004). Towards understanding assessments of the Big Five: Multitrait-multimethod analyses of convergent and discriminant validity across measurement occasion and type of observer. *Journal of Personality, 72*(4), 845–876. <https://doi.org/10.1111/j.0022-3506.2004.00282.x>
- Bleidorn, W., & Ostendorf, F. (2009). Ein Big Five-Inventar für Kinder und Jugendliche. [A Big Five inventory for children and adolescents: The German version of the Hierarchical

- Personality Inventory for Children (HiPIC)] *Diagnostica*, 55(3), 160–173.
<https://doi.org/10.1026/0012-1924.55.3.160>
- Booth, T., & Hughes, D.J. (2014). Exploratory structural equation modeling of personality data. *Assessment*, 21(3), 260–271. <https://doi.org/10.1177/1073191114528029>
- Caliendo, M., Fossen, F., & Kritikos, A.S. (2014). Personality characteristics and the decisions to become and stay self-employed. *Small Business Economist*, 42(4), 787–814.
<https://doi.org/10.1007/s11187-013-9514-8>
- Caspi, A., & Shiner, R.L. (2006). Personality development. In W. Damon, R. Lerner, & N. Eisenberg (Eds.), *Handbook of Child Psychology* (pp. 300–364). New York: Wiley.
- Caspi, A., Roberts, B.W., & Shiner, R.L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, 56(1), 453–484.
<https://doi.org/10.1146/annurev.psych.55.090902.141913>
- Chang, L., Connelly, B.S., & Geeza, A.A. (2012). Separating method factors and higher order traits of the Big Five: A meta-analytic multitrait-multimethod approach. *Journal of Personality and Social Psychology*, 102(2), 408–426. <https://doi.org/10.1037/a0025559>
- Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464–504.
<https://doi.org/10.1080/10705510701301834>
- Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255.
https://doi.org/10.1207/S15328007SEM0902_5
- Church, A.T., & Burke, P.J. (1994). Exploratory and confirmatory tests of the Big Five and Tellegen's three- and four-dimensional models. *Journal of Personality and Social Psychology*, 66(1), 93–114. <https://doi.org/10.1037/0022-3514.66.1.93>

- Costa, P.T., Jr., & McCrae, R.R. (1992). *NEO PI-R and NEO-FFI professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P.T., Jr., & McCrae, R.R. (1997). Longitudinal stability of adult personality. In R. Hogan, J.A. Johnson, S.R. Briggs, (Eds.), *Handbook of personality psychology* (pp. 269–290). San Diego, CA, US: Academic Press. <https://doi.org/10.1016/B978-012134645-4/50012-3>
- Costa, P.T., Jr., & McCrae, R.R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment*, *64*(1), 21–50.
- DeFruyt, F., Mervielde, I., Hoekstra, H.A., & Rolland, J.-P. (2000). Assessing adolescent's personality with the NEO-PI-R. *Assessment*, *7*(4), 329–345.
- DePauw, S.S.W., Mervielde, I., & Van Leeuwen, K.G. (2009). How are traits related to problem behavior in preschoolers? Similarities and contrasts between temperament and personality. *Journal of Abnormal Child Psychology*, *37*(3), 309–325. <https://doi.org/10.1007/s10802-008-9290-0>
- DeYoung, C.G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology*, *91*(6), 1138–1151. <https://doi.org/10.1037/0022-3514.91.6.1138>
- Fuchs, G., & Brunner, M. (2014). *Kompetenzentwicklung und Schulqualität an Brandenburger Grundschulen: Abschlussbericht der KEGS-Studie [Development of competencies in primary school: Final report of the KEGS study]*. Berlin: ISQ.
- Gerlitz, Y., & Schupp, J. (2005). *Assessment of Big Five personality characteristics in the SOEP*. German Institute of Economic Research, Berlin: DIW Research Notes 4.

- Guenole, N., & Brown, A. (2014). The consequences of ignoring measurement invariance for path coefficients in structural equation models. *Frontiers in Psychology, 5*, 980.
<https://doi.org/10.3389/fpsyg.2014.00980>
- Hahn, E., Gottschling, J., & Spinath, F.M. (2012). Short measurements of personality – Validity and reliability of the GSOEP Big Five Inventory (BFI-S). *Journal of Research in Personality, 46*(3), 355–359. <https://doi.org/10.1016/j.jrp.2012.03.008>
- Hildebrandt, A., Lüdtke, O., Robitzsch, A., Sommer, C., & Wilhelm, O. (2016). Exploring factor model parameters across continuous variables with local structural equation models. *Multivariate Behavioral Research, 51*(2-3), 257–258.
<https://doi.org/10.1080/00273171.2016.1142856>
- Hildebrandt, A., Wilhelm, O., & Robitzsch, A. (2009). Complementary and competing factor analytic approaches for the investigation of measurement invariance. *Review of Psychology, 16*(2), 87–102.
- Hu, L., & Bentler, P.M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*(4), 424–453.
<https://doi.org/10.1037/1082-989X.3.4.424>
- John, O.P., Caspi, A., Robins, R.W., Moffitt, T.E., & Stouthamer-Loeber, M. (1994). The “little five”: Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development, 65*, 160–178.
- John, O.P., Donahue, E.M., & Kentle, R.L. (1991). *The Big Five Inventory--Versions 4a and 54*. Berkeley, CA: University of California, Berkeley.
- John, O.P., Naumann, L.P., & Soto, C.J. (2008). Paradigm shift to the integrative Big Five Taxonomy. History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality*. (pp. 114–158). New York: Guilford.

- 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*. (pp. 102–139). New York: Guilford Press.
- Lang, F.R. (2005). *Erfassung des kognitiven Leistungspotenzials und der "Big Five" mit Computer-Assisted-Personal-Interviewing (CAPI): zur Reliabilität und Validität zweier ultrakurzer Tests und des BFI-S. [Assessment of cognitive capabilities and the Big Five with Computer-Assisted Personal Interviewing (CAPI): Reliability and validity]*. German Institute of Economic Research, Berlin: DIW Research Notes 9.
- Lang, F.R., John, D., Lüdtke, O., Schupp, J., & Wagner, G.G. (2011). Short assessment of the Big Five: Robust across survey methods except telephone interviewing. *Behavior Research Methods*, 43(2), 548–567. <https://doi.org/10.3758/s13428-011-0066-z>
- Little, T.D. (2013). *Longitudinal structural equation modeling*. New York, London: Guilford Press.
- Lucas, R.E., & Donnellan, M.B. (2011). Personality development across the life span: Longitudinal analyses with a national sample from Germany. *Journal of Personality and Social Psychology*, 101(4), 847–861. <https://doi.org/10.1037/a0024298>
- Lüdtke, O., Roberts, B.W., Trautwein, U., & Nagy, G. (2011). A random walk down university avenue: Life paths, life events, and personality trait change at the transition to university life. *Journal of Personality and Social Psychology*, 101(3), 620–637. <https://doi.org/10.1037/a0023743>
- Marsh, H.W. (1986). Negative item bias in rating scales for preadolescent children: A cognitive-developmental phenomena. *Developmental Psychology*, 22(1), 37–49.
- Marsh, H.W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A.J.S., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory

structural equation modeling. *Psychological Assessment*, 22(3), 471–491.

<https://doi.org/10.1037/a0019227>

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

McCrae, R.R., & Costa, P.T., Jr., (2008). The five-factor theory of personality. In O.P. John, R.W. Robins, & L.A. Pervin (Eds.), *Handbook of personality* (pp. 159–181). New York: Guilford.

McCrae, R.R., Costa, P.T., Jr., Terracciano, A., Parker, W.D., Mills, C.J., De Fruyt, F., & Mervielde, I. (2002). Personality trait development from age 12 to age 18: Longitudinal, cross-sectional and cross-cultural analyses. *Journal of Personality and Social Psychology*, 83(6), 1456–1468. <https://doi.org/10.1037//0022-3514.83.6.1456>

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

Measelle, J.R., John, O.P., Ablow, J.C., Cowan, P.A., & Cowan, C.P. (2005). Can children provide coherent, stable, and valid self-reports on the Big Five dimensions? A longitudinal study from ages 5 to 7. *Journal of Personality and Social Psychology*, 89(1), 90–106. <https://doi.org/10.1037/0022-3514.89.1.90>

Milfont, T.L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3(1), 111–121.

Muthén, L.K., & Muthén, B.O. (1998-2015). *Mplus user's guide. Seventh edition*. Los Angeles, CA: Muthén & Muthén.

Poropat, A.E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322–338.

- Putnick, D.L., & Bornstein, M.H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rammstedt, B., & Farmer, R.F. (2013). The impact of acquiescence on the evaluation of personality structure. *Psychological Assessment, 25*(4), 1137–1145. <https://doi.org/10.1037/a0033323>
- Roberts, B.W., & DelVecchio, W.F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*(1), 3–25.
- Roberts, B.W., Walton, K.E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1–25. <https://doi.org/10.1037/0033-2909.132.1.1>
- Rothbart, M.K., Ahadi, S.A., & Evans, D.E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology, 78*(1), 122–135. <https://doi.org/10.1037//0022-3514.78.1.122>
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23–74.
- Sirois, F.M., & Hirsch, J.K. (2015). Big Five traits, affect balance and health behaviors: A self-regulation resource perspective. *Personality and Individual Differences, 87*, 59–64. <https://doi.org/10.1016/j.paid.2015.07.031>
- Socio-Economic Panel (SOEP), data for years 1984-2014, version 31, SOEP, 2016, 10.5684/soep.v31.1.

- Soto, C.J., John, O.P., Gosling, S.D., & Potter, J. (2008). The developmental psychometrics of Big Five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology, 94*(4), 718–737.
<https://doi.org/10.1037/0022-3514.94.4.718>
- Specht, J., Egloff, B., & Schmukle, S.C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology, 101*(4), 862–882.
<https://doi.org/10.1037/a0024950>
- Steenkamp, J.-B.E.M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*(1), 78–107.
<https://doi.org/10.1086/209528>
- van der Linden, D., te Nijenhuis, J., & Bakker, A.B. (2010). The general factor of personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality, 44*(3), 315–327. <https://doi.org/10.1016/j.jrp.2010.03.003>
- Vassend, O., & Skrandal, A. (1997). Validation of the NEO Personality Inventory and the five-factor model. Can findings from exploratory and confirmatory factor analysis be reconciled? *European Journal of Personality, 11*, 147–166.
- Wagner, G.G., Frick, J.R., & Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) – Scope, evolution and enhancements. *Schmollers Jahrbuch, 127*(1), 139–169.
- Widaman, K.F., & Reise, S.P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K.J. Bryant, M. Windle, & S.G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). Washington DC: American Psychological Association.

Zentner, M., & Bates, J.E. (2008). Child temperament: An integrative review of concepts, research programs, and measures. *European Journal of Developmental Science*, 2(1/2), 7–37.

Ziegler, M., Poropat, A., & Mell, J. (2014). Does the length of a questionnaire matter? *Journal of Individual Differences*, 35(4), 250–261. <https://doi.org/10.1027/1614-0001/a000147>

Tables

Table 1

Measurement Invariance of Multi-Group CFA Models from Ages 11 to 84

Model	χ^2	df	<i>p</i>	CFI	Δ CFI	RMSEA [CI]	Δ RMSEA	SRMR	Δ SRMR	AIC	$\Delta\chi^2$ (Δ df) [compared to model x]
Model 1: configural	722.50	201	<.001	.981		.035[.033; .038]		.016		524924.57	
Model 2: metric	1006.59	342	<.001	.975	.006	.031[.029; .033]	.004	.024	.008	524995.81	287.51(141)***[1]
Model 3: scalar	2044.54	375	<.001	.939	.036	.046[.044; .048]	.015	.032	.008	526140.97	1258.00(33)***[2]
Model 3a: Item C2r Intercept free Age-Group 17-30	1772.52	374	<.001	.948	.027	.042[.040; .044]	.013	.030	.006	525823.34	910.78(32)***[2]
Model 3b: Item O2 Intercept free Age-Group 17-30	1654.48	373	<.001	.953	.022	.041[.039; .041]	.010	.029	.005	525686.56	764.77(31)***[2]
Model 3e: Item O4 Intercept free Age-Group 11-14	1575.34	372	<.001	.956	.019	.039[.037; .041]	.008	.029	.005	525595.79	667.25(30)***[2]
Model 3d: Item A2 Intercept free Age-Group 11-14	1504.25	371	<.001	.958	.017	.038[.036; .040]	.007	.028	.004	525513.24	585.05(29)***[2]
Model 3e: Item C2r Intercept free Age-Group 61-84	1436.98	370	<.001	.961	.014	.037[.035; .039]	.006	.028	.004	525436.77	501.01(28)***[2]
Model 3f: Item A3 Intercept free Age-Group 17-30	1380.92	369	<.001	.963	.012	.036[.034; .038]	.005	.027	.003	252373.49	431.53(27)***[2]
Age-Group 31-60	1297.40	368	<.001	.966	.009	.035[.033; .037]	.004	.027	.003	525277.74	329.03(26)***[2]
Model 4: Structural (with partial scalar intercepts)	1582.97	413	<.001	.957	.009	.037[.035; .039]	.002	.039	.012	525531.88	284.17(45)***[3f]

Note. Age groups 11-14 ($n=547$), 17-30 ($n=1507$), 31-60 ($n=4753$), 61-84 ($n=3104$). In models 3b-3f, item intercept equality constraints are relaxed in addition to previous relaxed intercepts. CI=Confidence interval (confidence level=.90)

Table 2

Item Labels and Intercepts (Model 3f Table 1)

Item Labels	Intercepts per age group			
	11-14	17-30	31-60	61-84
C1 ...does a thorough job	6.15 ⁱ			
C2r ...tends to be lazy	5.56	4.84	5.56	5.86
C3 ...does things efficiently	5.79 ⁱ			
A1r ...is sometimes rude to others	5.24 ⁱ			
A2 ...has a forgiving nature	6.06	5.53	5.53	5.53
A3 ...is considerate and kind to almost everyone	5.81	6.19	6.05	5.81
E1 ...is talkative	5.52 ⁱ			
E2 ...is outgoing, sociable	5.13 ⁱ			
E3r ...is reserved	3.69 ⁱ			
N1 ...worries a lot	4.24 ⁱ			
N2 ...gets nervous easily	3.65 ⁱ			
N3r ...is relaxed, handles stress well	3.36 ⁱ			
O1 ...is original, comes up with new ideas	4.56 ⁱ			
O2 ...values artistic, aesthetic experiences	4.37	3.83	4.37	4.37
O3 ...has an active imagination	4.80 ⁱ			
O4 ...is curious	4.76	5.49	5.49	5.49

Note. Group specific intercepts are reported for freed intercepts.

r=reverse-coded item. C=Conscientiousness, A=Agreeableness, E=Extraversion, N=Neuroticism, O=Openness to experience.

ⁱinvariant across groups

Table 3

Loadings of BFI-S Personality Items to Big Five Factors: Results of the Partial Scalar Invariance Model (Model 3f Table 1)

	λ		λ		λ		λ		λ
C		A		E		N		O	
C1	.742***	A1r	1.024***	E1	.981***	N1	.885***	O1	.882***
C2r	.680***	A2	.439***	E2	1.088***	N2	1.241***	O2	.985***
C3	.687***	A3	.750***	E3r	.932***	N3r	.725***	O3	1.081***
A1r	-.250***	E2	.063**	A2	.245***	C1	-.007	O4	.675***
E1	.195***	E3r	-.619***	A3	.185***	C3	-.035	C2r	-.243***
E3r	.002	N3r	-.206***	C2r	.082*	O1	.144	C3	.189***
N1	.247***	O1	-.340***	O1	.138***	O2	.365***	A1r	-.304***
O1	.281***	O2	.196***			O3	.315**	E1	-.168
O2	-.105	O4	.053*			O4	.027	E2	-.099
O3	-.043							E3r	-.275
O4	.220***							N3r	-.382***

Note. Unstandardized estimates are reported because equality constraints are based on unstandardized parameters. The standardized solution showed no loading greater than .3 of an item to a theoretically unintended factor in any age group except one loading from E3r to agreeableness across all age groups (Range of standardized loadings from E3r to A: $\lambda = -.404$ to $-.342$). C=Conscientiousness, A=Agreeableness, E=Extraversion, N=Neuroticism, O=Openness to Experience; r=reverse-coded item.

*** $p < .001$, ** $p < .01$, * $p < .05$.