Age differences in five personality domains across the life span

Allemand, M; Zimprich, D; Hendriks, A A J

Postprint available at:
http://www.zora.uzh.ch

Posted at the Zurich Open Repository and Archive, University of Zurich.
http://www.zora.uzh.ch

Originally published at:
Age differences in five personality domains across the life span

Abstract

The present study addresses the issue of age differences in five personality domains across the lifespan in a cross-sectional study. In contrast to most previous studies, we followed a methodologically more rigorous approach to warrant that age-related differences in personality structure and mean-level can be meaningfully compared. We used data on 50 items of the Five-Factor Personality Inventory (FFPI) available from a study in a large and representative Dutch sample (N = 2494, age range: 16-91 years) conducted in 1996 for the purpose of establishing norms for the FFPI. After having established strict measurement invariance, we tested for factor covariances to be equal across age groups, and found structural continuity of personality. Additionally, factor variances were shown to be equal across age groups. A number of age differences in the mean-level of the five personality domains emerged. Specifically, older adults were, on average, more agreeable and, especially, more conscientious than middle-aged and younger adults. Findings from our study suggest that both continuity and change may mark personality over the course of life.
Age Differences in Five Personality Domains across the Lifespan

Mathias Allemand, Daniel Zimprich
Department of Psychology, University of Zurich
&
A. A. Jolijn Hendriks
Groningen Institute for Educational Research, University of Groningen
Abstract

The present study addresses the issue of age differences in five personality domains across the lifespan in a cross-sectional study. In contrast to most previous studies, we followed a methodologically more rigorous approach to warrant that age-related differences in personality structure and mean-level can be meaningfully compared. We used data on 50 items of the Five-Factor Personality Inventory (FFPI) available from a study in a large and representative Dutch sample ($N = 2494$, age range: 16-91 years) conducted in 1996 for the purpose of establishing norms for the FFPI. After having established strict measurement invariance, we tested for factor covariances to be equal across age groups, and found structural continuity of personality. Additionally, factor variances were shown to be equal across age groups. A number of age differences in the mean-level of the five personality domains emerged. Specifically, older adults were, on average, more agreeable and, especially, more conscientious than middle-aged and younger adults. Findings from our study suggest that both continuity and change may mark personality over the course of life.

*Keywords:* Personality domains; personality change; measurement invariance; age differences; lifespan development
Introduction

Five broad domains—neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness—have been proposed to summarize individual differences in human personality traits (Digman, 1990; John, 1990). These five domains of personality have been consistently identified across numerous samples utilizing a variety of measurement instruments (John & Srivastava, 1999). Also, good self-other agreement has been found between self-ratings of personality and ratings made by peers and other observers as well as appropriate levels of cross-situational consistency (Funder & Colvin, 1997). In addition, these personality domains were found to demonstrate high levels of test-retest stability over time and age in terms of maintaining rank-order continuity (Fraley & Roberts, 2005; Roberts & DelVecchio, 2000). It has been argued that the five domains have biological foundations (Pickering & Gray, 1999), are highly heritable (Bouchard & Loehlin, 2001; Johnson, McGue, & Krueger, 2005), are found across a number of cultures (McCrae, 2001), and have analogues in our closest nonhuman relatives, chimpanzees (King & Figueredo, 1997). The common perspective of trait theories such as the Big Five (Goldberg, 1990) or the Five-Factor Model (McCrae & John, 1992), thus, is that personality traits are relatively enduring patterns of thoughts, feelings, and behaviors, which are expected to remain stable over time and are consistent across situations.

Consequently, the expectation is that there would be few developmental changes in personality across the adult lifespan. However, there is both cross-sectional and longitudinal evidence for age differences and systematic age-related changes in personality traits across the adult lifespan (e.g., De Fruyt, Van Leuven, Bagby, Rolland, & Rouillon, 2006; Helson, Jones, & Kwan, 2002; McCrae et al., 1999, 2005; Mroczek & Spiro, 2003; Roberts, Walton, & Viechtbauer, 2006; Small, Hertzog, Hultsch, & Dixon, 2003; Srivastava, John, Gosling, & Potter, 2003; Terracciano, McCrae, Brant, & Costa, 2005). These findings are in line with a lifespan developmental approach that is built on the core assumption that development is not
completed at a particular point in life, but extends over the entire lifespan (Baltes, Lindenberger, & Staudinger, 2006). The lifespan development perspective asserts that people are open systems and that they exhibit continuity, but also change in personality throughout the lifespan as a result of complex interactions between biological and socio-cultural influences, and the developing person (Baltes et al., 2006). It, thus, emphasizes the plasticity of psychological functioning across the lifespan, highlighting the possibility for changes to happen even in midlife and old age (e.g., Costa, Herbst, McCrae, & Siegler, 2000; Mroczek & Spiro, 2003; Small et al., 2003; Terracciano et al., 2005). Depending on the individual, similar processes may contribute to both continuity and change in personality in the form of genetic influences on developmental processes, responses to environmental circumstances, observational learning, learning generalizations, and learning from others’ descriptions of ourselves (for comprehensive reviews, see Caspi & Roberts, 2001; Roberts & Caspi, 2003).

In an attempt to reconcile previous findings, Roberts and Pomerantz (2004) asserted that the issue of age differences or changes in personality across the lifespan may be structured along multiple methodological perspectives. There are at least five different types of continuity and change: (1) Structural, (2) mean-level, (3) rank-order, (4) ipsative and (5) coherence (for details see Caspi & Roberts, 2001; Martin & Zimprich, 2005). While the emphasis is on continuity or change, which, strictly speaking, would require longitudinal data, structural and mean-level continuity and change may also be examined cross-sectionally, conditional on the assumption that cohort effects or interactions with cohort effects do not play a major role. Although longitudinal studies are generally preferable, large, well-designed cross-sectional studies provide useful estimates of age-related changes in mean levels of, e.g., personality traits (cf. Miyazaki & Raudenbush, 2000).

Mean-level change refers to changes in the average personality trait level of a population and is empirically assessed by examining group means over time or across different groups. This aspect of change is thought to result from maturational or historical
processes shared by a population. A number of cross-sectional and longitudinal studies have examined mean-level change of personality traits in adolescence, young adulthood, midlife, and old age (e.g., Costa et al., 2000; Helson et al., 2002; Helson & Soto, 2005; McCrae et al., 1999, 2000; Roberts, Caspi, & Moffitt, 2001; Roberts et al., 2006; Small et al., 2003; Terracciano et al., 2005). For example, Srivastava et al. (2003) studied age differences in personality in a large cross-sectional sample of more than 130,000 internet users with an age range from 21 to 60 years. They found that conscientiousness and agreeableness increased throughout early and middle adulthood. Srivastava and colleagues' findings were similar to those reported by McCrae et al.'s (1999, 2000) multi-national studies with a total sample size of over 12,000 adults, where, across cultures, the median correlations of age with neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness scales were -0.17, -0.21, -0.08, 0.09, and 0.23, respectively. With respect to age differences or age-related changes in personality domains during the last decades of life, it seems that neuroticism decreases across older adulthood (Mroczek & Spiro, 2003) and may show some increase very late in life (Small et al., 2003). Recently, Weiss et al. (2005) examined cross-sectional age trends in personality among Medicare patients aged 65 to 100 and found that age was positively related to agreeableness ($r = 0.18$). A recent literature review by Roberts, Robins, Caspi, and Trzesniewski (2003) summarized previous cross-sectional and longitudinal data on mean-level continuity and change of the five personality domains. They concluded that, on average, people become more agreeable and more conscientious through midlife and old age. In addition, people show decreases in neuroticism across all age periods and small increases in openness to experience in the early stages of young adulthood and little change thereafter (see also McCrae & Costa, 2003). Results for extraversion are less consistent, unless one organizes the literature around two components underlying this domain, namely social dominance and social vitality (cf. Helson & Kwan, 2000): People, on average, increase in measures of social dominance and decrease on measures of social vitality with age (Roberts et
al., 2003, 2006). However, the distinction into those two components is not beyond controversy (Costa & McCrae, 2006). With respect to higher-level extraversion, McCrae and Costa (2003) concluded that, between ages 30 and 80, there is a small decline in this broad personality domain. To summarize, previous findings indicate that mean levels of personality domains continue to change during adulthood into old age, especially with respect to neuroticism, agreeableness and conscientiousness.

However, few studies have examined age differences or age changes in personality after having established measurement invariance. Measurement invariance (MI) means that indicators, e.g., items of a personality inventory, of an underlying latent construct, e.g., neuroticism, mean the same things to members of different groups such as age groups. In other words, MI implies that measurement bias with respect to groups is absent, such that factors or latent variables are equally scaled (Meredith, 1993; Meredith & Horn, 2001). In essence, measurement invariance thus reflects an issue of construct validity. Strictly speaking, age-related differences in personality structure and personality means across groups can be meaningfully studied only if its measurement is unbiased across groups. In many studies it has been implicitly assumed that the measures utilized to assess personality be invariant, an assumption that, if it goes untested or is only partially tested, may lead to an over- or underestimation of age-related differences in personality. Note, however, that MI represents one part of structural continuity. Assessing structural continuity encompasses two related, but distinct parts, which are briefly discussed below: (1) Assessing invariance of the measurement part of the model as a necessary condition for (2) assessing invariance of the structural part of the model.

(1) Invariance of the measurement model. The measurement part of the model specifies the relations between latent variables (personality factors) and their manifest indicators, e.g., questionnaire scores. Equivalence of these relations across groups has been labeled measurement invariance (MI) in the psychometric literature (cf. Bollen, 1989; Meredith,
As Horn and McArdle (1992) have defined it, MI refers to “whether or not, under different conditions of observing and studying phenomena, measurement operations yield measures of the same attribute” (p. 117). Assuming that one has applied the same multiple items (or scales) measuring different personality constructs in different groups defined by a selection variable, e.g., age, MI may be evaluated by examining invariance in factor loadings, latent intercepts, and residual variances by means of confirmatory factor analyses of personality questionnaires across groups. MI is an issue of degree, which, borrowing from Meredith’s (1993) terminology, ranges from configural invariance to strict measurement invariance (see method section). Examining different degrees of MI is accomplished by employing multiple-group confirmatory factor models with increasingly severe across-group restrictions on parameters (cf. Allemand, Zimprich, & Hertzog, 2007; Martin & Zimprich, 2005; Zimprich, Allemand, & Hornung, 2006).

Several studies have examined different degrees of MI of personality measures across age-groups and over time by utilizing confirmatory factor analyses (CFI). For example, Small et al. (2003) were able to establish weak measurement invariance of the NEO-PI across a 6-year longitudinal period in older adults. Likewise, Morizot and Le Blanc (2003) found partial weak measurement invariance, i.e., the majority of factor loadings remained invariant, of personality scales across two age groups and across time. Recently, Allemand et al. (2007) demonstrated that, both cross-sectionally and longitudinally, strict measurement invariance of the NEO-FFI held in two adult samples (445 participants aged 42-46 and 420 participants aged 60-64) followed across four years. As another approach of examining whether the personality factor structure is invariant across ages, other researchers reported congruence coefficients (e.g., Allik, Laidra, Realo, & Pullmann, 2004; Lang, Lüdtke, & Asendorpf, 2001; Srivastava et al., 2003). The typical procedure is to perform exploratory factor analyses (EFA) within age groups and extract the expected number of factors. In order to compare the factorial structure across age groups, Procrustes rotation (cf. McCrae, Zonderman, Costa, &
Bond, 1996) is used, and subsequently, amount of congruence is established. Using this approach, Srivastava et al. (2003) found invariance in the pattern of factor loadings of the Big Five Inventory (BFI) across four age groups. They reported an average congruence coefficient across age groups of .99, reflecting a high degree of similarity of factors. Likewise, Lang et al. (2001) found an invariant factor structure of the German version of the BFI across three age cohort groups, i.e., young, middle-aged, and old adults. To summarize, there is some, but limited evidence regarding MI of personality measures across age.

(2) Invariance of the structural model. The structural part of the model specifies the associations among a set of latent variables (personality factors). If these associations are stable across groups, structural continuity holds. Structural continuity builds upon MI, because, at least, weak MI has to be established in order to render comparisons of covariances among personality factors meaningful (Horn & McArdle, 1992; Meredith, 1993; Meredith & Horn, 2001). Empirically, it involves investigating the similarity of variances and covariances among personality factors across groups. The importance of structural continuity stems from the fact that, e.g., decreasing associations among personality factors into adulthood would imply an unfolding of personality, i.e., personality becoming more diversified and less unitary. By contrast, should associations for instance increase into old age, personality would become more unified (cf. Baltes, Cornelius, Spiro, Nesselroade, & Willis, 1980). From a developmental perspective, such a differentiation or dedifferentiation appears of great relevance because it represents a prerequisite for assuming that different or the same causes underlie the development in different personality factors (see Hofer & Sliwinski, 2001).

Only a few studies have examined the degree of structural continuity as based on weak MI across age groups or over time using multiple-group CFA. Small et al. (2003), for instance, found personality factor covariances to be equal longitudinally in older adults, indicating high structural continuity over time. Recently, based on strict MI, Allemand et al. (2007) demonstrated invariant covariation patterns cross-group and cross-time in a sample of
middle-aged and older adults, indicating that the five-factor personality covariance structure was highly stable. Robins, Fraley, Roberts, and Trzesniewski (2001) reported highly stable interrelations among the NEO-FFI personality factors in young adulthood over a 4-year period. Reviewing previous research, Costa and McCrae (1997) concluded that cross-sectional personality structure seems to be invariant at different ages. To summarize, previous findings suggest relatively high levels of structural continuity of the five personality domains across age groups and over time. However, not always has weak MI been established as a necessary condition.

In the present study we also examined an additional type of continuity and change: Continuity of divergence (e.g., Allemand et al., 2007; Martin & Zimprich, 2005). Continuity of divergence refers to the fact that the amount of interindividual differences in personality factors might increase, decrease, or remain stable across age. Empirically, this type of continuity can be examined by comparing personality factor variances across age groups. An increase or decrease of personality factor variances would indicate that the amount of change is different for different persons. If we borrow from the literature on cognitive development, different amounts of individual differences in personality domains might be indicative of the variables governing change and development. Horn (1988) has argued that relatively homogenous developmental trajectories might characterize a more biologically-driven developmental process, whereas increasing variances might denote changes triggered by external influences that are socially driven. Based on this line of reasoning one would expect a variance increase with age of those factors being more socially driven, e.g., agreeableness. By contrast, the amount of individual differences should remain relatively stable across age for those personality traits that are more genetically based, e.g., extraversion (cf. Johnson et al., 2005).

To our knowledge, only two studies have rigorously tested for continuity of divergence in the five personality domains. Small et al. (2003) found that the Big Five personality factor
variances were equal across a 6-year period in a sample of older adults, implying continuity of divergence over time. In addition, Allemand et al. (2007) reported that the openness for experience variance in middle-aged participants was significantly larger than in older participants at two measurement occasions. That is, the sample of older participants was more homogeneous than the one of middle-aged participants with respect to the propensity to be creative, complex, and open to new ideas.

The Present Study

In the present study, we set out to understand age differences in personality across the lifespan by examining the five personality domains. A cross-sectional design was used to study how personality domains differ by age. Moreover, in the present study we used data on a large and representative sample with a continuous age distribution to test continuity and change from young adulthood into old age, i.e., age 16 to 91. This allowed us to take a lifespan perspective on the five dimensions of personality and to examine and clarify age differences and change from young adulthood into midlife and late life. The specific aims were the following: (1) To test assumptions about increasing levels of measurement invariance (MI), to study structural continuity and continuity of divergence among the five personality domains across age groups, (2) to investigate age differences in the factor means across the lifespan after having established at least strong MI.

Method

Sample

We used data from the normative sample of the Five-Factor Personality Inventory (FFPI) gathered in the Netherlands (Hendriks, Hofstee, & De Raad, 1999a, 1999b). The sample comprised 2494 participants (1367 males, 1127 females) who completed the FFPI and other questionnaires in the context of an ongoing survey, which aims to monitor savings and investment behavior in the Netherlands (Hendriks et al., 1999a, 1999b). For this survey, panel members periodically complete various questionnaires in return for having free use of a
personal computer. Questionnaires were downloaded, answered, and uploaded after completion by participants. At the time (1996) the data were collected, participants’ average age was 46.4 years (SD = 15.4), ranging from 16 to 91 years. Participants highest level of education was: 3.7% primary school, 35.1% secondary school, 30.3% high school, and 30.9% university (Hendriks et al., 1999a, 1999b).

For the present study, we divided the sample into six age groups: (1) 16-29 years (M = 21.9 years, N = 316), which was chosen to be the reference group, (2) 30-39 years (M = 34.9 years, N = 519), (3) 40-49 years (M = 44.1 years, N = 652), (4) 50-59 years (M = 54.2 years, N = 441), (5) 60-69 years (M = 64.1 years, N = 364), (6) 70+ years (M = 74.6 years, N = 202). Although the youngest and the oldest age groups comprised smaller sample sizes than the other groups, all groups were sufficiently large (N > 200).

**Instruments**

We used a representative selection of items of the Five-Factor Personality Inventory (FFPI; Hendriks, 1997; Hendriks et al., 1999a, 1999b; Hendriks, Hofste, & De Raad, 2002), which assesses a person’s position on extraversion, agreeableness, conscientiousness, emotional stability (or, conversely, neuroticism), and autonomy (weakly related to openness). The FFPI has been developed in the tradition of the lexical approach to personality description. FFPI items consist of brief behavioral descriptions (e.g., “engages in discussions,” “loves order and regularity”) as an alternative to trait adjective rating scales. Ratings are made on a 5-point Likert scale ranging from 1 (not at all applicable) to 5 (entirely applicable) with higher scores indicating more pronounced values on the five respective personality dimensions. The rationale for the development of concrete behavioral sentence items instead of trait adjectives was that adjectives are abstract terms and therefore more difficult for a broad range of educational levels. Indeed, Hendriks (1997) found that many trait adjectives are unfamiliar to the less highly educated respondents, whereas behavioral translations of trait adjectives are not. Several studies provided support for the FFPI being a
reliable and valid instrument, which shows more than adequate psychometric properties and cross-cultural generalizability (e.g., Barelds & Luteijn, 2002; Hendriks et al., 1999a, 1999b; Hendriks, Kuyper, Offringa, & Van der Werf, in press; Hendriks et al., 2003; Perugini & Ercolani, 1998).

Note that autonomy bears only limited resemblance to openness to experience of the NEO-PI-R (e.g., Costa, Yang, & McCrae, 1998; De Fruyt, McCrae, Szirmák, & Nagy, 2004; Perugini & Ercolani, 1998; for a full discussion of similarities and differences, see Hendriks, 1997, pp. 79-81). FFPI-autonomy appears to capture critical reflection and making one’s own choices (Hendriks et al., 1999b). In line with this, De Fruyt et al. (2004) suggested that this domain might be interpreted as a dominance factor. Dominance, in turn, forms part of extraversion, which might be characterized as combining sociability, i.e., the care about social interactions, with an active and adventurous engagement with the world (Helson & Kwan, 2000). By contrast, Hmel and Pincus (2002) demonstrated that autonomy appears to be closest to self-governance, which may be considered as sharing some overlap with conscientiousness.

Note that since we utilized the FFPI, empirically distinguishing between social dominance and social vitality parts of extraversion (cf. Helson & Kwan, 2000) was impossible.

From the 100 items of the FFPI, we selected 50 (10 items for each personality domain). Item selection was necessary due to the following reason: The factor model underlying the FFPI is non-congeneric, that is, individual items may load on more than one common factor while, at the same time, factors are orthogonal (Hendriks et al., 1999a, 1999b, 2002). By contrast, in the present study our aim was to arrive at a congeneric model of (possibly correlated) personality factors, that is, individual items may load on one factor only. Hence, for each personality domain we chose those 10 items with the lowest cross-loadings on other personality factors. As will be outlined below, the necessity to have items that share one underlying dimension is also given by utilizing item parceling to form manifest indicators of personality (see below).
Overview of Statistical Analyses

Multiple-group confirmatory factor analysis (CFA) including means was utilized in order to assess measurement invariance, structural and mean-level continuity and change across age (cf. Bollen, 1989). Models are described in more detail below. First, however, we will present two features common to all models, namely parceling and the way models were parameterized.

Parceling. Rather than using individual items as indicators of the five latent factors, we chose to use parcels each made up of 3-4 items (cf. Bandalos & Finney, 2001; Little, Cunningham, Shahar, & Widaman, 2002). A parcel may be defined as an aggregate-level indicator comprised of the sum (or average) of several single items that can reasonably be considered to have the same underlying dimension. Before constructing parcels, we tested this unidimensionality of the items being parceled as a prerequisite (Bandalos & Finney, 2001). Subsequently, parcels were built according to the Item-to-Construct Balance technique (Little et al., 2002, p. 166). Briefly, the three items with the highest loadings were selected to anchor the three parcels of each personality factor. Subsequently, the three items with the next highest item-to-construct loadings were added to the anchor parcels in an inverted order. This procedure was repeated until all items had been assigned to a parcel. As a result, for each personality factor two parcels consisting of three items each and one parcel consisting of four items each were built. Note that compared to individual items as indicators of latent constructs, parceling offers some potential benefits (cf. Little et al., 2002). Because parcels are more likely to be normally distributed than single items, the assumptions underlying maximum likelihood parameter estimation are more easily met. Moreover, the resulting reduction in the complexity of measurement models achieved by parceling may lead to more precise and stable parameter estimates.  

Parameterization. A common approach to parameterize latent constructs (factors) in confirmatory factor models is to identify factor variances and means by setting the loading of
one manifest reference variable to one and the intercept of this reference variable to zero. Then, the factor is scaled like the reference variable and the factor mean is equal to the intercept of the reference variable. One potential problem of this approach in the context of multiple-group models is that by fixing one factor loading to one it is implicitly assumed that this parameter is invariant across different groups. Moreover, this approach confounds group differences in factor means and group differences in the intercepts of the manifest indicators used as reference variables (cf. Meredith & Horn, 2001). Therefore, we utilized an alternative parameterization. Let \( x_g = (x_{g1}, x_{g2}, \ldots, x_{gq})^T \) denote the vector of manifest indicators in group \( g \) \((g = 1, \ldots, G)\). A common factor model in group \( g \) may then be written as (cf. Bollen, 1989)

\[
\begin{align*}
    x_g &= \tau_g + \Lambda_g \xi_g + \delta_g ,
\end{align*}
\]

where \( \tau_g \) is a vector of latent intercepts, \( \Lambda_g \) is a matrix of factor loadings, \( \xi_g \) is a vector of common factors, and \( \delta_g \) is a vector of residuals. Define \( \Phi_g \), the (co-)variance matrix of the common factors in group \( g \), \( \Theta_g \), the (co-)variance matrix of residuals in group \( g \), \( \mu_g \), the means of the manifest indicators in group \( g \), and \( \kappa_g \), the means of the common factors in group \( g \). A model for the covariances \( \Sigma_g \) among the manifest indicators in group \( g \) then is

\[
\begin{align*}
    \Sigma_g &= \Lambda_g \Phi_g \Lambda_g^T + \Theta_g ,
\end{align*}
\]

and a model for the means \( \mu_g \) of the manifest indicators in group \( g \) is

\[
\begin{align*}
    \mu_g &= \tau_g + \Lambda_g \kappa_g .
\end{align*}
\]

For identification we chose \( \text{diag}(\Phi) = I \), i.e., the variances of the latent variables \( \xi \) were set to 1 in all groups. In addition, factor means were constrained to be zero in all groups, i.e., \( \kappa = 0 \), and latent intercepts of all manifest indicators were estimated instead. These constraints were later relaxed depending on the model specified and its identification status. Specifically, after having established strict measurement invariance (see below) across age groups, those constraints were retained for the youngest age group, the reference group, whereas for the
other age groups factor means and factor variances were freely estimated. Note that the estimated factor means and variances then represent relative values that have to be interpreted in comparison with the reference group.

Measurement invariance. To examine measurement invariance (MI), different degrees of measurement invariance of the five domain scales (extraversion, agreeableness, conscientiousness, emotional stability, and autonomy) were imposed by constraining parameters to be equal across age groups, as is detailed below. Meredith (1993; see also Meredith & Horn, 2001) distinguished between four increasingly restrictive levels of measurement invariance: (1) Configural invariance, (2) weak measurement invariance, (3) strong measurement invariance, and (4) strict measurement invariance (for a full discussion of MI, see Horn & McArdle, 1992; Meredith, 1993; Meredith & Horn, 2001).

Configural invariance entails that the number of factors and the according salient and non-salient loadings are equal across age groups, which ensures that the dimensionality of the measured constructs is equivalent. Weak measurement invariance might be assumed when it can be demonstrated that pattern matrices be fully invariant across age groups ($\Lambda_g = \Lambda$). On a conceptual level, weak measurement invariance ensures that the relationships between the indicators (manifest variables) and the concepts (latent variables)—represented by the magnitude of the factor loadings—are equivalent across age groups. Strong measurement invariance, a more stringent form of measurement invariance, involves consideration of the means on both the manifest and the latent variables (scalar invariance). In this study, the hypothesis of strong measurement invariance was tested by fitting models with constraints on the measurement intercepts—that is, the intercepts in the regression models that relate each indicator to the latent concept. Strong measurement invariance requires that pattern matrices and latent intercepts of the manifest indicators are invariant across age groups ($\Lambda_g = \Lambda, \tau_g = \tau$). Strict measurement invariance involves additional constraints, namely that measurement uniquenesses, i.e., residual variances, are equivalent across groups
\[ \Lambda_g = \Lambda, \tau_g = \tau, \Theta_g = \Theta \]. Note that if weak MI holds, comparisons of factor (co-) variances, and if strong measurement invariance holds, comparisons of factor means across groups are rendered meaningful (cf. Meredith, 1993; Meredith & Horn, 2001). Strict MI implies that all of the differences in means, covariances, and variances of the observed indicators are genuine differences between groups not attributable to differential psychometric properties of the measurement instrument.

**Examining different types of continuity and change.** After having established strict measurement invariance, factor covariances were compared between age groups to examine structural continuity. In order to test for statistically significant differences, equality constraints were imposed on the factor covariances across age groups. Note that we compared factor covariances, because, by comparing correlations one implicitly assumes that factor variances are also equal. Additionally, in order to test for continuity of divergence, i.e., continuity in the amount of interindividual variability in the five personality domains across age, we tested a model in which factor variances were constrained to be equal across age groups. Eventually, to assess mean-level change in the five personality domains, factor means were compared, with the youngest age group functioning as the reference group.

All analyses were conducted using MPLUS version 3.0 (Muthén & Muthén, 2004). The goodness-of-fit of the models was evaluated using the \( \chi^2 \)-test. As additional criteria for absolute model fit the Comparative Fit Index (CFI), and the Root Mean Square Error of Approximation (RMSEA) are reported. Values of the CFI above .90 denote a well-fitting model, whereas for the RMSEA values less than .06 indicate an acceptable model fit (cf. Browne & Cudeck, 1993). In comparing the relative fit of nested models, we used the \( \chi^2 \)-difference test. Due to its dependency on sample size, the \( \chi^2 \)-difference test provides rather high power for large sample sizes. We therefore complemented it by calculating 90% RMSEA confidence intervals for the models estimated (MacCallum, Browne, & Sugawara,
Since the RMSEA is virtually independent of sample size, the comparison of RMSEA confidence intervals provides an effective, alternative method of assessing relative model fit of nested models. As a measure of effect size for mean differences, we report Cohen’s $d$ (Cohen, 1988, p. 20). Given the large sample size in this study, the $\alpha$-level was set to 1% in order to evaluate statistical significance.

**Results**

**Measurement Invariance**

The multiple-group confirmatory factor analyses (CFAs) started with an unconstrained model, that is, a *configural invariance* model with five factors of personality without any parameter constraints across age groups (Model 1). Factor variances were fixed to 1 and factor means were fixed to 0 in order to scale the latent variables. Model 1 achieved a good fit (see Table 1). Next, in Model 2, factor loadings were constrained to be equal across age groups, while factor variances were freely estimated in all age groups apart from the youngest group, i.e., the reference group of those aged 16-29 years. Model 2 also evinced a good fit (see Table 1). In comparison to Model 1, Model 2 did not represent a statistically significant reduction in relative fit. Also, as indexed by the overlap of the RMSEA 90% confidence intervals, there was no difference in fit. From these results, one might conclude that *weak measurement invariance* holds across the age groups with respect to the five personality domains.

In Model 3, the additional constraint of equal latent intercepts of the manifest indicators, implying strong measurement invariance, was tested. Factor means were freely estimated in all age groups except from the reference group, i.e., those aged 16-29 years. As Table 1 shows, Model 3 achieved an acceptable fit. Compared to Model 2, Model 3 represented a statistically significant reduction in relative fit. However, the 90% CIs of the RMSEAs did exhibit overlap, indicating that, according to the RMSEA, model fit was indistinguishable.
Hence, we concluded that strong measurement invariance holds across the age groups with respect to the five personality dimensions.

Subsequently, in Model 4, strict measurement invariance was tested, i.e., residual variances were constrained to be equal across age groups. Model 4 evinced an acceptable fit as well (see Table 1). Compared to Model 3, there was a statistically significant loss of fit as indexed by the $\chi^2$-difference test. The overlapping RMSEA 90% confidence intervals, however, suggested that the difference in model fit was not of practical importance, indicating that the hypothesis of strict measurement invariance should not be rejected. Model 4, the model of strict measurement invariance, seemed to adequately capture our data.

Taken together, the measurement properties of the instrument used to operationalize the five personality domains might be considered invariant across the six age groups.

Subsequently, invariance of the interrelations among the personality factors and of the variances of the personality factors across the different age groups was investigated.

*Structural Continuity and Continuity of Divergence*

In order to test for structural continuity across age groups, first, factor covariances were constrained to be equal across age groups (Model 5; see Table 1). Model 5 achieved an acceptable fit. In comparison to Model 4, Model 5 did not represent a statistically significant loss in fit. In line with this, the RMSEA confidence intervals showed considerable overlap. Therefore, equal factor covariances could be assumed in all age groups, implying structural continuity.

Subsequently, in Model 6, factor variances were constrained to be equal across age groups. Model 6 evinced an acceptable fit (see Table 1). Albeit, compared to Model 5, there was a statistically significant reduction in model fit, the RMSEA 90% CIs suggested that, from a practical point of view, model fit was indistinguishable. We therefore concluded that individual differences in the five personality domains were equally pronounced in all age groups. As a consequence, not only are factor covariances equal across age groups, but—due
to equal factor variances—factor correlations were age-invariant, too. Latent factor correlations are reported in Table 2. Note that the highest factor correlation emerged between extraversion and autonomy \((r = 0.448)\), whereas agreeableness and autonomy were unrelated \((r = 0.001)\). The median correlation among factors was \(r = 0.269\), indicating medium-sized associations between the five personality domains (cf. Cohen, 1988, p. 80). Parameter estimates based on Model 6 and the amounts of explained variance in the manifest indicators for all six age groups are shown in Table 3.

**Mean-Level Change**

In order to examine mean-level continuity and change in the five personality domains, factor means were compared. Although, in principle, in order to compare factor means across groups it is sufficient to establish strong measurement invariance, the equality of factor variances across age groups demonstrated to hold by Model 6 has a convenient advantage: Because factor variances are all equal to one in all age groups, factor means, which represent differences to the youngest age group, can be interpreted directly as effect sizes (Cohen’s \(d\)).

To determine whether two age groups differ significantly from each other with respect to factor means on the 5%-level, we calculated 84% confidence intervals (CIs) for independent group means. Goldstein and Healy (1995) demonstrated that 95% CIs can overlap substantially and yet two means will be significantly different at the 5% level. They showed that if the standard errors of two independent means are approximately equal, then non-overlapping 84% CIs are equivalent to a \(t\)-test of these means at the .05 level (see also Tyron, 2001). Table 4 and Figure 1 show the age differences in factor means, using the youngest age group, i.e., 16-29 years, as the reference group having factor means of zero, that is, factor means in the other groups were scaled as deviations from the reference group.

Table 4 and Figure 1 are to be read as follows: If the 84% confidence interval (CI) of a factor mean in one age group overlaps with the 84% CI of the corresponding factor mean in another age group, factor means are not significantly different at the 5%-level. In turn, if the
84% CI of a factor mean in one age group does not overlap with the 84% CI of the corresponding factor mean in another age group, factor means should be considered as being significantly different at the 5%-level. For example, the 84% CI of the mean of agreeableness in those aged 60-69 years ranges from 0.162 to 0.444. The 84% CI estimate of the mean of agreeableness of those aged 16-29 years ranges -0.149 to 0.150. Hence, those aged 60-69 years are, on average, more agreeable than those aged 16-29 years (Cohen’s $d = 0.303$). As another example, the 84% CI of the mean of emotional stability in those aged 70+ years ranges from -0.173 to 0.166. The 84% CI of the mean of emotional stability in those aged 30-39 years ranges from 0.035 to 0.266. Hence, those aged 70+ years do not differ significantly from those aged 30-39 years with respect to emotional stability (Cohen’s $d = 0.146$).

Both in terms of statistical significance and effect sizes, the picture that emerged with respect to means in personality domains may be described as follows (see Table 4 and Figure 1): (a) Extraversion generally showed a decrease across age groups, implying that older adults are, on average, less extraverted than younger adults. However, in the oldest group, extraversion appeared to be somewhat more pronounced again. Note, however, that these tendencies in neither case were statistically significant. Concordantly, effect sizes were small. (b) Agreeableness showed an increase across age groups, implying that the elderly were, on average, more agreeable than younger adults. However, in the oldest age group, agreeableness appeared to be somewhat less pronounced than one would have expected if agreeableness showed a monotonic increase across age groups. Again, effect sizes were relatively small. (c) Conscientiousness exhibited a monotonic and comparatively pronounced increase across age groups, which is reflected in a number of statistically significant differences and large effect sizes, especially the differences between the youngest and the two oldest age groups. On average, conscientiousness in individuals from the oldest age group was almost one standard deviation above the factor mean of the youngest age group, amounting to an effect size of $d = .85$. (d) Emotional stability did not follow a clear age trajectory; instead, factor means seemed
to fluctuate without any apparent tendency to increase or decrease across age. Mean differences among age groups were all in the range of small effects. (e) Finally, autonomy showed a decrease across age groups, the decrease being nearly significant in the two oldest age groups. This implies that the elderly were, on average, somewhat less autonomous than younger adults. Effect sizes were small, however.

In summary, our results with respect to age differences in the five factor means across the adult lifespan show a clear trend for increases in agreeableness and, most notably, conscientiousness. Older participants in the present study were, on average, more agreeable and much more conscientious than younger adults. Small, non-significant age differences were found for extraversion and autonomy, showing a slight decrease across age. By contrast, emotional stability exhibited inconsistent age differences. For reasons of completeness and comparability with other studies, we also calculated correlations between age and the five latent personality factors. These age correlations were $r = -.07$ (extraversion), $r = .10$ (agreeableness), $r = .24$ (conscientiousness), $r = -.02$ (emotional stability), and $r = -.07$ (autonomy). Apart from the correlation between age and emotional stability, all correlations were statistically significant ($p < .01$).

Discussion

The purpose of our study was to examine age differences in the five personality domains across the adult lifespan in a large and representative sample. Before analyzing age differences in the structure and mean-levels of personality, we examined measurement invariance (MI) of the personality instrument. We found that the criteria for strict measurement invariance across age groups were met (e.g., Meredith & Horn, 2001). Note, however, that our inferences about measurement invariance are tempered by the fact that we did not evaluate invariance across the individual 50 items selected from the FFPI. By using parcels, we specified a less complex measurement model than others (e.g., Small et al., 2003), which probably contributed to the feasibility of finding strict measurement invariance in the
present study. However, taking into account the severity of restrictions to the models and the large sample size, both implying high statistical power (cf. MacCallum et al., 1996), and the fact that—with a fully-fledged five-factor measurement model—quite a number of parameters are involved, the finding of strict MI with respect to the FFPI across six age groups still appears remarkable. Also, strict measurement invariance with respect to age allows for extrapolations with respect to other selection variables, e.g., health, because it almost certainly implies weak measurement invariance for all selection variables correlated to age (Lubke, Dolan, Kelderman, & Mellenbergh, 2003).

Next, based on strict measurement invariance, the covariance patterns of the five personality factors were compared across age groups. We found no indication of any practically important age differences in associations among personality domains. According to this finding, across the six age groups structural continuity of the five personality factors holds, which gives support to Costa and McCrae’s (1997) assertion that, after adolescence, the structure of personality is constant across age. Although, in general, highly stable interrelations among the five personality domains across age have been reported, both cross-sectionally (Allemand et al., 2007; Costa & McCrae, 1997; Lang et al., 2001; Srivastava et al., 2003) and longitudinally (Allemand et al., 2007; Caspi & Roberts, 2001; Robins et al., 2001; Small et al., 2003), the present result extends previous findings because a more rigorous methodology was applied, a larger age range was covered, and complete structural continuity on the factorial level emerged. From a substantive perspective, structural continuity as observed in the present study—even with mean-level age differences in personality being present—is consistent with the idea that normative changes in the majority of personality domains are modest in magnitude and might act to preserve, not alter, the structure of personality (Terracciano et al., 2005). Thus, we found no indication of cross-sectional differentiation nor dedifferentiation of personality traits across the lifespan. This contrasts, for example, with findings on cognitive changes across the lifespan, where mean age changes or
age differences are much more pronounced, and where, repeatedly, slightly increasing associations among ability factors across age have been reported, a phenomenon termed dedifferentiation (e.g., Babcock, Laguna, & Roesch, 1997; Hertzog & Bleckley, 2001). Subsequently, in addition to factor covariances, factor variances were constrained to be equal across age, which did not lead to any practically important decrement of model fit. This finding implies that the amount of interindividual variability in the five personality domains was constant across the six age groups, implying continuity of divergence across age—an issue that has rarely been addressed in previous research on age changes in personality, albeit it has long been acknowledged as a fundamental characteristic of the study of aging (Dannefer, 1988; Nesselroade, 1991). Note that the combination of equality of factor or “true” variances and strict measurement invariance, that is, equality of factor loadings and “error” variances, implies equal reliabilities of the manifest indicators across the six age groups (cf. Bollen, 1989). Due to the cross-sectional nature of the data analyzed in the present study strong conclusions about personality variance continuity appear unwarranted, but as a result it matches previous longitudinal findings (Allemand et al., 2007; Small et al., 2003). There are reasons why one might have expected increasing interindividual differences with age. For example, the combined effects of individuals’ unique experiences over more years might have produced increasing differences among them. Also, genetically based differences would have had more time to be expressed and to cause individuals to diverge. Moreover, older people, somewhat freed of societal constraints, might be more likely to choose their own courses of action. However, it appears as if, with respect to personality, there is no “age-related heterogeneity” (Dannefer, 1988), at least not until the age of 75, the mean age of the oldest group included in the present study. It remains open, then, what happens to personality variability during the “fourth age” (Baltes & Smith, 1999). One ramification of age-invariant factor covariances and age-invariant factor variances is that correlations among the five personality factors were also equal across the six age
groups, apart from sampling error. Note that this is a stronger finding than structural continuity alone, because it implies that structural continuity of the five personality factors is scale invariant, that is, insensitive to a change in scaling of the personality factors (Cudeck, 1989). Regarding the size of factor interrelations, in contrast to our approach, most published studies on personality domains are based on orthogonal, non-congeneric factor analyses that produced uncorrelated factors (e.g., McCrae et al., 1996; Goldberg, 1992). By contrast, we chose a congeneric, confirmatory approach that led to correlated factors. Digman (1997) reanalyzed a data set from Goldberg, and the factor correlations he reported are, in general, weaker than in the present sample. One reason for this might be that, although both Goldberg’s Big Five Markers and the FFPI are based on the psycho-lexical approach, the former is composed of trait adjective scales while the latter consists of brief behavioral descriptions and they differ with respect to the abstractness of the indicators, i.e., items (cf. Goldberg, 1992; Hendriks et al., 1999a, 1999b). Also, Digman (1997) conducted an analysis based on already extracted personality factors, and the method with which these factors were obtained remains unclear, although this may have a considerable impact on results (Fabrigar, Wegener, MacCallum, & Strahan, 1999).²¹ In comparison to the Dutch normative sample of the NEO-FFI (Hoekstra, Ormel, & De Fruyt, 1996), the correlations between extraversion and agreeableness and extraversion and autonomy (openness) were elevated, as were the correlations between agreeableness and conscientiousness, and between emotional stability and autonomy (openness). In turn, the correlations between extraversion and conscientiousness and between emotional stability and conscientiousness were lower in the present sample. Consistent with the assumption that autonomy reflects, in part, a dominance factor (cf. De Fruyt et al., 2004), and shares some overlap with conscientiousness (cf. Hmel & Pincus, 2002), we found extraversion and conscientiousness being the strongest correlates of autonomy.
Pertaining to mean-level age differences, we found a small decrease in extraversion with age; however, these tendencies were not statistically significant nor of relevant effect size. Roberts et al. (2003, 2006) pointed out that previous studies also did not demonstrate a clear pattern of mean-level age differences or change in extraversion unless this domain is differentiated into two distinct components, social dominance and social vitality (cf. Helson & Kwan, 2003). Such a distinction was not possible using FFPI data, hence testing this hypothesis was beyond the scope of the present investigation. By contrast, clear age increases were found in agreeableness and conscientiousness. Similar to previous research in terms of effect sizes, our findings add to converging evidence that agreeableness and, even more so, conscientiousness increase across the lifespan (e.g., Lang et al., 2001; McCrae, 2000; Srivastava et al., 2003). The mixed age trend of emotional stability found in the present study fits into previously reported findings on the inconsistent nature of age differences in neuroticism (e.g., Helson & Kwan, 2000; McCrae et al., 1999; Roberts et al., 2003, 2006). Finally, for autonomy comparable cross-sectional results were reported, for instance, by McCrae et al. (1999) and Smith and Baltes (1999), who found negative age trends in openness to experience. Additionally, longitudinal evidence for a decline in openness in older age has, recently, been documented by Roberts et al. (2006). However, one has to keep in mind that autonomy should not be equated to openness to experience. It would, thus, be informative to further examine age differences and age-related changes in autonomy in future studies.

Age-related mean-level differences in agreeableness, conscientiousness, and emotional stability, although, for the latter domain, not found in the present study, may be described as an increase in personality maturity, in the sense of becoming emotionally more predictable and more attuned to social demands (e.g., Caspi, Roberts, & Shiner, 2005; Helson & Wink, 1987; Roberts & Caspi, 2003; Roberts & Wood, 2006). Irrespective of its causal pathway, the pattern of personality changes enhancing maturity across adulthood contributes to everyday life running smoothly, to maintaining or augmenting subjective well-being, life success, and
longevity. For example, lifespan studies have shown that individuals who score high on traits of conscientiousness/constraint and positive emotionality live longer (Danner, Snowdon, & Friesen, 2001; Friedman et al., 1995). By contrast, individuals high in traits opposite to agreeableness, e.g., anger and hostility, are at greatest risk of disease, e.g., cardiovascular illness (Miller, Smith, Turner, Guijarro, & Hallet, 1996).

Recently, Staudinger and Kunzman (2005) have argued that this specific configuration of personality changes might be the result of successful coping with normative developmental tasks and challenges of adulthood and, thus, increased adjustment. These normatively-triggered changes in personality domains may help in preparing people for dealing with normative developmental tasks. Regarding non-normative age-related personality changes, by contrast, Baltes and colleagues (Baltes et al., 2006) emphasized the potential importance of individual-specific life events in old age as a cause of development. Just as normative life events, such as retirement or loss of a spouse in old age (e.g., Field & Millsap, 1991), non-normative life events in midlife, e.g., changes in jobs or marital status, can alter personality (e.g., Costa et al., 2000).

A limitation of the present study is the fact that we used a cross-sectional design to make inferences about personality differences and, consequently cannot differentiate between developmental and cohort effects. Differential sampling by age and cohort differences are both potential sources of confounds. That is, differences, e.g., in mean-levels of conscientiousness, found across age groups can be attributed, in part, to the culture, climate or historical context that an individual was born into and lived through (Hofer & Sliwinski, 2001). However, the comparison of findings from cross-sectional and longitudinal studies can provide insight. As previously noted, our results with respect to MI and structural continuity are comparable with the findings from longitudinal studies (e.g., Allemand et al., 2007; Small et al., 2003). In addition, our results concerning mean-level change agree with the broad trends among cross-sectional and longitudinal studies reviewed by Roberts et al. (2003, 2006).
Also, similar age differences in the Big Five personality traits were found in a variety of cultures (McCrae, 2001; McCrae et al., 1999) and in comparable chimpanzee personality dimensions (King, Weiss, & Farmer, 2005). However, longitudinal studies using the FFPI are needed in order to examine the longitudinal trajectory of the personality domain autonomy across the lifespan.

Another limitation of the present study refers to the fact that the youngest age group spans from ages 16 to 29. Although it is widely recognized that every phase in life is somehow sensitive for development, Bornstein (1989) noted that theory and data signify that some periods in life may be more critical than others. Indeed, studies on personality trait development (e.g., Roberts & DelVecchio, 2000; Roberts et al., 2006; Srivastava, et al., 2003) have shown that earlier periods in life are marked by a higher degree of change in personality traits compared to midlife and old age. Some clarification would be provided, for example, by dividing age groups into smaller age spans in young adulthood, which, however, was not possible in the present study because otherwise age groups would have become too small for meaningful multiple-groups analyses.

To close, in the present paper extensive and systematic age-comparative analyses on a representative subsample of items of the FFPI were conducted, as an operationalization of the Big Five personality domain. As a prerequisite, we first established strict measurement invariance to hold across age. Our findings demonstrate that the instrument behaved equivalently across six age groups and yielded an invariant factorial structure across age. Subsequently, we focused on a thorough description of several aspects of personality continuity and change across the lifespan. The broader developmental picture that emerged from the present study is one of structural continuity and one of mean-level age differences in the five broad personality domains across the adult lifespan, highlighting the possibility for changes to happen even in midlife and old age.
References


### Appendix

<table>
<thead>
<tr>
<th>Parcels</th>
<th>Item numbers from the original FFPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXTRA1</td>
<td>46, 61, 91, 96</td>
</tr>
<tr>
<td>EXTRA2</td>
<td>31, 41, 86</td>
</tr>
<tr>
<td>EXTRA3</td>
<td>66, 71, 76</td>
</tr>
<tr>
<td>AGRE1</td>
<td>57, 67, 77, 97</td>
</tr>
<tr>
<td>AGRE2</td>
<td>17, 82, 92</td>
</tr>
<tr>
<td>AGRE3</td>
<td>27, 47, 72</td>
</tr>
<tr>
<td>CONS1</td>
<td>3, 33, 73, 88</td>
</tr>
<tr>
<td>CONS2</td>
<td>8, 63, 78</td>
</tr>
<tr>
<td>CONS3</td>
<td>13, 38, 98</td>
</tr>
<tr>
<td>EMOS1</td>
<td>29, 39, 59, 64</td>
</tr>
<tr>
<td>EMOS2</td>
<td>14, 34, 74</td>
</tr>
<tr>
<td>EMOS3</td>
<td>79, 84, 99</td>
</tr>
<tr>
<td>AUTO1</td>
<td>20, 25, 40, 60</td>
</tr>
<tr>
<td>AUTO2</td>
<td>30, 65, 80</td>
</tr>
<tr>
<td>AUTO3</td>
<td>35, 50, 100</td>
</tr>
</tbody>
</table>

*Note.* Parcels of extraversion: EXTRA1 to EXTRA3, parcels of agreeableness: AGRE1 to AGRE3, parcels of conscientiousness: CONS1 to CONS3, parcels of emotional stability: EMOS1 to EMOS3, and parcels of autonomy: AUTO1 to AUTO3. Information on item content can be obtained from the third author at [jolijn.hendriks@planet.nl](mailto:jolijn.hendriks@planet.nl).
Author Note

Mathias Allemand and Daniel Zimprich, Department of Psychology, Gerontopsychology, University of Zurich; A. A. Jolijn Hendriks, Faculty of Behavioral and Social Sciences, Groningen Institute for Educational Research (GION), University of Groningen.

The order of the first two authors is strictly alphabetical, both contributed equally.

We would like to thank Brent W. Roberts, Toni C. Antonucci, and three anonymous reviewers for their helpful suggestions on earlier versions of the article.

A. A. J. Hendriks is the first author of the Five-Factor Personality Inventory (FFPI). The FFPI is commercially published in The Netherlands.

Correspondence concerning this article should be addressed to Mathias Allemand, Department of Psychology, Gerontopsychology, University of Zurich, Binzmühlestrasse 14/24, CH-8050 Zurich, Switzerland. E-mail: m.allemand@psychologie.uzh.ch
Footnotes

1) Note that albeit in applications of the (full) FFPI acquiescence variance is removed, our modelling approach that relied on item parcels of part of the FFPI items did not meaningfully allow to do so.

2) An anonymous reviewer pointed out that the factor intercorrelations might be consistent with an underlying higher-order factor of socially desirable traits (Digman, 1997). However, empirically testing this hypothesis was beyond the scope of the present investigation, although the equality of factor correlations across age groups would imply that such a second order factor would also be stable across age groups.

3) Note that in applications of the (full) FFPI the uncorrelated factor scores are used rather than the correlated scale scores (e.g., Hendriks et al., 1999b).
Table 1

*Fit Indices for Multiple Group Models (N = 2494)*

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$\Delta \chi^2$</th>
<th>$\Delta df$</th>
<th>CFI</th>
<th>RMSEA</th>
<th>RMSEA 90% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>928.58*</td>
<td>480</td>
<td>—</td>
<td>—</td>
<td>0.975</td>
<td>0.047</td>
<td>0.043; 0.052</td>
</tr>
<tr>
<td>Model 2</td>
<td>988.10*</td>
<td>530</td>
<td>59.52</td>
<td>50</td>
<td>0.974</td>
<td>0.046</td>
<td>0.041; 0.050</td>
</tr>
<tr>
<td>Model 3</td>
<td>1235.44*</td>
<td>580</td>
<td>247.34*</td>
<td>50</td>
<td>0.963</td>
<td>0.052</td>
<td>0.048; 0.056</td>
</tr>
<tr>
<td>Model 4</td>
<td>1417.05*</td>
<td>655</td>
<td>181.61*</td>
<td>75</td>
<td>0.957</td>
<td>0.053</td>
<td>0.049; 0.057</td>
</tr>
<tr>
<td>Model 5</td>
<td>1487.95*</td>
<td>705</td>
<td>70.90</td>
<td>50</td>
<td>0.956</td>
<td>0.052</td>
<td>0.048; 0.055</td>
</tr>
<tr>
<td>Model 6</td>
<td>1565.38*</td>
<td>730</td>
<td>77.43*</td>
<td>25</td>
<td>0.953</td>
<td>0.052</td>
<td>0.049; 0.056</td>
</tr>
</tbody>
</table>

*Note.* $\chi^2$ = Chi-square, df = Degrees of Freedom, $\Delta \chi^2$ = Chi-square Difference, $\Delta df$ = Degrees of Freedom Difference, CFI = Comparative Fit Index, RMSEA = Root Mean Square Error of Approximation, RMSEA 90% CI = 90% Confidence Interval (CI) of RMSEA; Model 1 = model of configural invariance, Model 2 = model of weak measurement invariance, Model 3 = model of strong measurement invariance, Model 4 = model of strict measurement invariance, Model 5 = Model 4 plus the additional constraints of equal factor covariances across age groups, Model 6 = Model 5 plus the additional constraints of equal factor variances across age groups.  

*p < .01.
Table 2

*Latent Interfactor Correlations (N = 2494)*

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Extraversion</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Agreeableness</td>
<td>0.371</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Conscientiousness</td>
<td>0.139</td>
<td>0.270</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Emotional stability</td>
<td>0.425</td>
<td>0.193</td>
<td>0.267</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5. Autonomy</td>
<td>0.448</td>
<td>0.001</td>
<td>0.099</td>
<td>0.367</td>
<td>-</td>
</tr>
</tbody>
</table>

*Note.* All estimated interfactor correlations are statistically significant ($p < .01$), except for the correlation between agreeableness and autonomy (0.001).
Table 3

*Parameter Estimates of Model 6 (N = 2494)*

<table>
<thead>
<tr>
<th>Parcels</th>
<th>Factor loading</th>
<th>Latent intercept</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXTRA1</td>
<td>1.831</td>
<td>10.822</td>
<td>0.708</td>
</tr>
<tr>
<td>EXTRA2</td>
<td>1.880</td>
<td>10.865</td>
<td>0.685</td>
</tr>
<tr>
<td>EXTRA3</td>
<td>1.758</td>
<td>14.169</td>
<td>0.648</td>
</tr>
<tr>
<td>AGRE1</td>
<td>0.961</td>
<td>8.159</td>
<td>0.461</td>
</tr>
<tr>
<td>AGRE2</td>
<td>0.880</td>
<td>8.110</td>
<td>0.559</td>
</tr>
<tr>
<td>AGRE3</td>
<td>0.849</td>
<td>10.628</td>
<td>0.534</td>
</tr>
<tr>
<td>CONS1</td>
<td>1.508</td>
<td>10.411</td>
<td>0.501</td>
</tr>
<tr>
<td>CONS2</td>
<td>1.523</td>
<td>11.102</td>
<td>0.626</td>
</tr>
<tr>
<td>CONS3</td>
<td>1.502</td>
<td>15.504</td>
<td>0.620</td>
</tr>
<tr>
<td>EMOS1</td>
<td>1.746</td>
<td>10.850</td>
<td>0.669</td>
</tr>
<tr>
<td>EMOS2</td>
<td>1.709</td>
<td>11.785</td>
<td>0.698</td>
</tr>
<tr>
<td>EMOS3</td>
<td>1.695</td>
<td>15.904</td>
<td>0.688</td>
</tr>
<tr>
<td>AUTO1</td>
<td>2.024</td>
<td>6.147</td>
<td>0.749</td>
</tr>
<tr>
<td>AUTO2</td>
<td>2.142</td>
<td>6.664</td>
<td>0.746</td>
</tr>
<tr>
<td>AUTO3</td>
<td>1.658</td>
<td>9.127</td>
<td>0.667</td>
</tr>
<tr>
<td>Mean $R^2$</td>
<td>-</td>
<td>-</td>
<td>0.637</td>
</tr>
</tbody>
</table>

*Note.* Parcels of extraversion: EXTRA1 to EXTRA3, parcels of agreeableness: AGRE1 to AGRE3, parcels of conscientiousness: CONS1 to CONS3, parcels of emotional stability: EMOS1 to EMOS3, and parcels of autonomy: AUTO1 to AUTO3. Factor loadings are unstandardized. Minima and maxima of the explained variance in the manifest indicators are underscored.
Table 4

**Factor Means and 84% Confidence Intervals (CI) Based on Model 6 (N = 2494)**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Extraversion</th>
<th>Agreeableness</th>
<th>Conscientiousness</th>
<th>Emotional stability</th>
<th>Autonomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-29 years</td>
<td>M 0+</td>
<td>0.082</td>
<td>0.406</td>
<td>-0.140; 0.141</td>
<td>-0.021</td>
</tr>
<tr>
<td></td>
<td>84% CI</td>
<td>-0.140; 0.140</td>
<td>0.147; 0.147</td>
<td>-0.140; 0.141</td>
<td>-0.139; 0.138</td>
</tr>
<tr>
<td>30-39 years</td>
<td>M 0+</td>
<td>0.082</td>
<td>0.507</td>
<td>0.035; 0.266</td>
<td>-0.021</td>
</tr>
<tr>
<td></td>
<td>84% CI</td>
<td>-0.139; 0.093</td>
<td>0.526</td>
<td>-0.013; 0.200</td>
<td>-0.135; 0.093</td>
</tr>
<tr>
<td>40-49 years</td>
<td>M 0+</td>
<td>0.138</td>
<td>0.672</td>
<td>0.094</td>
<td>-0.048</td>
</tr>
<tr>
<td></td>
<td>84% CI</td>
<td>-0.256; -0.043</td>
<td>0.617</td>
<td>0.094</td>
<td>-0.154; 0.057</td>
</tr>
<tr>
<td>50-59 years</td>
<td>M 0+</td>
<td>0.237</td>
<td>0.815</td>
<td>-0.005</td>
<td>-0.048</td>
</tr>
<tr>
<td></td>
<td>84% CI</td>
<td>-0.309; -0.063</td>
<td>0.799</td>
<td>-0.029; 0.216</td>
<td>-0.165; 0.057</td>
</tr>
<tr>
<td>60-69 years</td>
<td>M 0+</td>
<td>0.303</td>
<td>0.846</td>
<td>-0.137; 0.128</td>
<td>-0.219</td>
</tr>
<tr>
<td></td>
<td>84% CI</td>
<td>-0.328; -0.063</td>
<td>0.952</td>
<td>-0.137; 0.128</td>
<td>-0.350; -0.088</td>
</tr>
<tr>
<td>70+ years</td>
<td>M 0+</td>
<td>0.284</td>
<td>0.997</td>
<td>-0.173; 0.166</td>
<td>-0.232</td>
</tr>
<tr>
<td></td>
<td>84% CI</td>
<td>-0.306; 0.034</td>
<td>0.696; 0.997</td>
<td>-0.173; 0.166</td>
<td>-0.390; -0.073</td>
</tr>
</tbody>
</table>

Note. + = fixed parameter. All estimated parameters are relatively scaled with the youngest age group being the reference group.
Figure Captions

Figure 1. Age differences in personality across the lifespan ($N = 2494$) based on Model 6. Age groups: (1) 16-29 years, (2) 30-39 years, (3) 40-49 years, (4) 50-59 years, (5) 60-69 years, (6) 70+ years. Personality domains: Extraversion (E), agreeableness (A), conscientiousness (C), emotional stability (ES), and autonomy (AU). Note that factor means, which represent differences to the youngest age group, can be interpreted directly as effect sizes, i.e., Cohen’s $d$ (see result section).