Empirical Paper

EUROPEAN JOURNAL OF
Personality

European Journal of Personality
0(0) 1–21

DOI: 10.1177/08902070211014054
journals.sagepub.com/home/ejop
SAGE

# Correlated personality change across time and age

**Gabriel Olaru[1] and Mathias Allemand[2]**

## Abstract

The goal of this study was to examine differential and correlated change in personality across the adult lifespan. Studying differential and correlated change can help understand whether intraindividual trait change trajectories deviate from the norm and how these trajectories are coupled with each other. We used data from two large longitudinal panel studies from the United States that covered a total age range of 20 to 95 years on the first measurement occasion. We used correlated factor models and bivariate latent change score models to examine the rank-order stability and correlations between change across three measurement waves covering 18 years ($N = 3250$) and four measurement waves covering 12 years ($N = 4145$). We examined the moderation effects of continuous age on these model parameters using local structural equation modeling. The results suggest that the test–retest correlations decrease with increasing time between measurements but are unaffected by participants' age. We found that change processes in Extraversion, Openness, Agreeableness, and Conscientiousness were strongly related, particularly in late adulthood. Correlated change patterns were highly stable across time intervals and similar to the initial cross-sectional Big Five correlations. We discuss potential mechanisms and implications for personality development research.

**Keywords**

personality, development, correlated change, rank-order stability, local structural equation modeling

Received 1 February 2021; Revised 19 March 2021; accepted 5 April 2021

## Introduction

Research on personality development suggests that personality change and stability can be evaluated from multiple perspectives (Caspi & Roberts, 2001; Costa et al., 2019; Roberts et al., 2008). Each perspective provides a unique view about the ways in which basic personality processes unfold over time. One perspective that has received much attention in the field of personality development is differential change. Differential change refers to the degree to which people maintain their standing on a given personality variable relative to others over time. The evidence for rank-order stability in the Big Five personality traits in adulthood appears to be strong and suggests relatively high persistence of individual differences over time (Roberts & DelVecchio, 2000; Terracciano et al., 2006; Wagner et al., 2019).

Another perspective that has received little attention in the field of personality development is correlated change (Allemand et al., 2007; Allemand & Martin, 2016). It refers to the question of whether and to what degree changes in personality are interconnected or affect each other. For example, as people become more conscientious over time, do they also become more agreeable and less neurotic? Or do personality traits develop independently over time? What is the degree of commonality in the change processes? Correlated change is theoretically important because it provides a unique view on personality development that complements the perspective of differential change (Allemand & Martin, 2016). Whereas differential change addresses the rank-order of change in a single variable, correlated change covers the amount of commonality in rank-orders of change across two or more variables. Furthermore, whereas differential change examines the correlation between static factors across measurement occasions, correlated change focuses on the correlations between dynamic change factors (Allemand

[1]Department of Developmental Psychology, Tilburg University, Netherlands
[2]Department of Psychology & URPP "Dynamics of Healthy Aging," University of Zurich, Switzerland

**Corresponding author:**
Gabriel Olaru, Department of Developmental Psychology, Tilburg University, Prof. Cobbenhagenlaan 225, 5037 DB Tilburg, Netherlands.
Email: golaru@mail.de

& Martin, 2016). Finally, correlated change has the great potential for a better understanding of the dynamic nature of personality, as it focuses on the commonality between different ways in which basic personality processes unfold over time (Allemand & Martin, 2016). Correlated change is also practically important because it has the potential to inform intervention programs aimed at self-improvement and volitional personality change (Hudson & Fraley, 2015; Stieger et al., 2020). The knowledge that personality change processes unfold together or in different ways could be informative for the development of intervention efforts to support and guide people who would like to modify or change personality traits. Despite the importance of this unique perspective, to date only limited research has been conducted on correlated personality change.

The present work thus sought to extend the literature on personality development in two ways. First, we examined differential and correlated personality change across different time intervals and developmental periods in two large longitudinal panel studies using the same measure of the Big Five personality traits. Second, we examined whether and to what degree time interval and continuous age moderate differential and correlated personality change across the adult lifespan.

### *Differential and correlated personality change*

Differential change provides an important perspective for the understanding of personality development (Caspi & Roberts, 2001; Costa et al., 2019; Roberts et al., 2008). Personality development through the lens of differential change refers to the degree to which the relative ordering of individuals on a given personality variable is maintained or changed over time. This aspect of change is most often assessed through test–retest correlations or stability coefficients of measurement occasions separated by a specified time interval. This perspective addresses the question whether people retain their standing on a variable relative to others over time.

If people were rather stable in a given personality variable over time or they do change, but in more or less the same way, one would expect high rank-order stability coefficients. This can occur when a normative event such as retirement impacts all individuals in the same way. In contrast, if people changed over time in different directions in relation to one another, one would expect low rank-order stability coefficients. This can occur when people are exposed to unique experiences or when the factors that influence personality are normative but people differ in their unique reactions and coping strategies to these events. Finally, it is also possible that measurement error or less reliable measurements produce low rank-order stability coefficients.

Correlated change is an important and complimentary view on personality development and provides information about the pattern and degree of commonality in the ways in which basic personality processes unfold over time (Allemand et al., 2007; Allemand & Martin, 2016). This perspective is essential for a better understanding of personality development and the exploration of underlying mechanisms that potentially shape development. It addresses the question whether change processes across variables are largely influenced by a few general, broadly acting mechanisms or by specific, narrowly acting mechanisms, each affecting only one single personality trait. If personality changes were rather general and shared similar causes, one would expect high correlations among intraindividual changes in different personality traits. That is, on an individual level, longitudinal changes in personality traits should be proportional to each other. On a group level, changes in personality traits should then be highly correlated, which may suggest that personality traits work together as a system to produce particular developmental trajectories and outcomes.

Broadly acting mechanisms, such as genetic factors, developmental processes, social roles, environments, and life events (Caspi & Roberts, 2001; Kandler et al., 2015; Roberts et al., 2008; Specht et al., 2014; Wrzus & Roberts, 2017), that operate simultaneously on multiple personality traits in the same way would result in a large degree of correlated personality change. Correlations between change trajectories might also indicate a bidirectionality or feedback loops in the change processes (Klimstra et al., 2013), with change in one trait affecting other trait domains—and vice versa. This can indicate transactional personality change processes (e.g. Roberts & Wood, 2006), which reinforce each other to achieve desired end states (e.g. *corresponsive principle*; Caspi et al., 2005; Roberts & Wood, 2006). From the perspective of self-regulated personality change (e.g. Hennecke et al., 2014), finding correlated change can also indicate that individuals try to change their trait levels in tandem in order to achieve their desired end states. Current findings on volitional or intervention-based personality change (Hudson & Fraley, 2015; Hudson et al., 2019; Roberts et al., 2017; Stieger et al., 2020) research suggest that individuals can actively affect their trait levels. The desire to change personality traits seems to be moderately correlated across trait domains (Hudson & Fraley, 2015), and self-regulated change or interventions may affect several traits simultaneously (Roberts et al., 2017). Correlated change might thus be observable when a group of individuals within the sample tries to change several trait-related habits or behaviors in tandem, whereas others maintain stability during this time. Whereas these different intraindividual trajectories will only be observable as moderate mean-level differences when examining absolute change, the perspective of differential and correlated change can help uncover such interindividual differences in coupled intraindividual development trajectories.

## *Previous work on differential and correlated personality change*

There is accumulating evidence for relatively high levels of rank-order stability over time from early adulthood to advanced old age, but this does not imply that there are no reliable individual differences in personality change (Borghuis et al., 2017; Chopik & Kitayama, 2018; Ferguson, 2010; Klimstra et al., 2009; Roberts & DelVecchio, 2000; Terracciano et al., 2006; Wagner et al., 2019). A conclusion based on previous work is that stability declines as a function of time interval. If traits are modified by experience, it would be reasonable to assume that the vicissitudes of human life would move people's trait scores in random ways, wandering further and further from initial levels. A recent meta-analysis estimated that the average observed value after an interval of 15 years would be approximately $r = .60$ (Anusic & Schimmack, 2016).

Another conclusion based on previous work is that stability generally increases during adolescence and young adulthood—an effect that is described as the *cumulative stability principle* (Caspi et al., 2005; Roberts & Wood, 2006). However, due to inconsistencies in the literature, it is not entirely clear whether and to what degree these stability correlations vary as a function of age after young adulthood. For example, Roberts and DelVecchio (2000) concluded that their findings indicate relatively high and increasing levels of differential stability across the lifespan with a peak of $r = .74$ around 60 years of age, whereas Ferguson (2010) found a relatively high stability ($r = .82$ to .94) from 30 to 80 years of age. Other studies with broad age ranges that also included advanced old age suggest that differential stability may decrease again after the peak in the sixth decade of life and thus reflects a *diversification* in old age (Lucas & Donnellan, 2011; Specht et al., 2011; Wagner et al., 2019; Wortman et al., 2012). In summary, available research suggests relatively high levels of rank-order stability over time in adulthood and a potential decline in old age.

Research on correlated change is prominent in the field of cognitive aging (Hülür et al., 2015; Martin & Zimprich, 2003; Mascherek & Zimprich, 2011; Sliwinski et al., 2003; Zimprich & Martin, 2002, 2010). However, it has received little attention in the field of personality development. To date, only a few studies have examined correlated personality change and their findings have been inconsistent (see Allemand & Martin, 2016 for a review). The first study examined correlated personality change in cohorts of middle-aged and older adults over a period of four years (Allemand et al., 2007). The results demonstrated considerable amounts of commonality between changes in personality traits with average absolute change correlations of $|r| = .37$ and .32 for middle-aged and older adults, respectively (Neuroticism was negatively correlated with the other traits). The correlated change patterns were similar to the initial trait correlations (middle age: $|r| = .26$; old age: $|r| = .32$). A follow-up study examined long-term correlated personality change across 12 years in the older age cohort (Allemand et al., 2008). Although a certain degree of commonality in the degree of personality change was observed, the patterns of correlated change painted a different picture with respect to Neuroticism. Whereas the initial study (Allemand et al., 2007) found that changes in Neuroticism were related to changes in the other traits with an average correlation of $r = -.44$, these correlations were close to zero in the follow-up study (Allemand et al., 2008). Across the other Big Five factors, the average change correlation was $r = .54$, whereas the average cross-sectional correlation was $r = .39$.

A study of 15-year change in a sample of young adults (Mund & Neyer, 2016) found an average absolute initial correlations between Neuroticism, Extraversion, Agreeableness, and Conscientiousness of $|r| = .32$ at the start of the study (again Neuroticism correlated negatively with the other traits). The 15-year change factors correlated on average by $|r| = .40$. In a sample of 81-year-old participants with a repeated measurement 6 years later (Mõttus et al., 2012), an initial average absolute correlation of $|r| = .22$ and absolute change correlation of $|r| = .17$ were found. Similar correlational patterns at the initial measurement and between change factors (initial $r = .17$; change $r = .13$) were observed in a large and representative sample ($N = 14{,}886$) that covered ages from 17 to 96 years with a repeated measurement after a 4-year time period (Klimstra et al., 2013). Correlated change was strongest for Extraversion and Openness ($r = .26$), as well as Agreeableness and Conscientiousness ($r = .25$). These two trait pairs also showed the highest cross-sectional level correlations in the study ($r = .37$ and $r = .32$, respectively). The somewhat smaller overall correlation coefficients compared to the other studies can be attributed to the use of scale scores instead of latent variable modeling to represent the personality traits. Finally, a study tested long-term correlated personality change over a 40-year time interval in a sample of 125 woman beginning at the age of 21 (Soto & John, 2012). The results, however, did not find significant correlations between personality change over 40 years, which may be due to the small sample size and the very long interval of time between measurement occasions.

To the best of our knowledge, only two studies have compared correlated personality change across age groups. First, although Allemand et al. (2007) did not explicitly test whether patterns of correlated change were equal across two distinct age groups, the average absolute change correlations of .37 and .32 for middle-aged and older adults were very similar. Second, the results of an explicit test of age moderation across 14 age groups covering an age range from 17 to 96 years indicated that the amount of correlated personality change was relatively stable from adolescence through middle adulthood and then increased in old age (Klimstra et al., 2013). The increase in correlated change was strongest for the Extraversion-Openness, Extraversion-Conscientiousness, and Openness-Conscientiousness pairs. Interestingly, these trait pairs also showed an increase in the cross-sectional level correlations across age. Klimstra et al. (2013) argued that the age effects in correlated change in late adulthood are related to broadly acting mechanisms, such as the dopaminergic system (DeYoung & Gray, 2009) affecting Openness and Extraversion, as well as developmental changes in cognitive abilities in old age (Craik & Bialystok, 2006).

In summary, available research suggests that there is some commonality in change between personality traits and that the degree of commonality between such changes may be similar to the initial trait commonalities. However, the few existing reports differed in many important ways (see Allemand & Martin, 2016). Apart from differences in measurement instruments, sample size, and modeling procedures (including the use of parceling techniques versus manifest scale scores as representations of the traits), we identified two critical differences amongst the studies. First, the studies differed with respect to the length of time between assessments. The time interval between repeated measurements ranged from 4 to 40 years across studies. The length of time between assessments has a known positive effect on mean-level change and a negative effect on rank-order stability, implying that larger normative and interindividual changes occur as more time passes between assessments (Roberts et al., 2006; Roberts & DelVeccio, 2000). Whether this effect also holds for correlated change has never been examined, as no study has yet examined correlated change as a function of time between measurement occasions.

Second, the studies differed with respect to the age of the participants in the samples. Some studies examined correlated change in narrow age ranges, while other studies focused on broader age ranges. Only the two aforementioned studies reported correlated change across age, but both used categorical age groups (Allemand et al., 2007; Klimstra et al., 2013). According to the cumulative stability principle (Caspi et al., 2005; Roberts & Wood, 2006), personality becomes increasingly more stable across adulthood. However, some studies also suggest that stability decreases again in old age (Lucas & Donnellan, 2011; Wagner et al., 2019; Wortman et al., 2012). How age, and in particular old age, affects correlated change is so far only addressed in the study by Klimstra et al. (2013), who found an increase in correlated change between Extraversion, Openness, and Conscientiousness in late adulthood, but otherwise relatively high stability. However, these findings were based on manifest scale scores as representations of the personality traits and categorical age groups to examine the moderating effect of age.

### *Local structural equation modeling*

To address the issue of categorical age variables, we made use of local structural equation modeling (LSEM; Briley et al., 2015; Hildebrandt et al., 2009, 2016; Olaru et al., 2019). LSEM is a procedure to examine moderation effects of continuous variables (in this case age) on all model parameters within a structural equation modeling (SEM) context. In the context of multigroup confirmatory factor analysis (MGCFA), continuous moderator variables are categorized (e.g. age groups) to achieve sufficient sample sizes for the model estimation, when the number of observations at each level (e.g. years of age) is too small. This artificial categorization is problematic, because participants close to the cutoffs can be more similar across groups than within groups. For instance, when forming age groups based on decades, 29 and 31 year olds will be assigned to different groups despite being more similar in age than 31- and 39-year-old participants. The allocation of cutoffs will inevitably influence the results (see Hildebrandt et al., 2009; MacCallum et al., 2002), and particularly broad or heterogeneous groups may mask underlying differences. Identifying nonlinear moderation effects or critical change points with a low number of groups can be very difficult.

LSEM overcomes this issue by estimating an SEM at every age value using sampling weights instead of categorical age groups for the model estimation—weighting each participant based on the distance to the target age value (i.e. higher weights for participants closer to the target age). The underlying rationale is that participants with smaller age differences are also more similar in the psychological constructs to be studied than participants with larger age differences, thus weighting participants accordingly. For estimating the model parameters for 50 years of age (i.e. "local"), the difference of participants' age from 50 years of age will be used to determine the sampling weights. Participants with a difference of 0 (i.e. 50 years old) will be fully weighted, and the other participants with partial weights that are decreasing with the distance to 50 years (e.g. 49 and 51 will receive higher weights than 48 and 52, which are weighted more than 47 and 53, etc.). The weighting function follows a Gaussian kernel function (see Hildebrandt et al., 2009, 2016; Olaru et al., 2019 for more details). The breadth of this weighting kernel can be modified to increase or decrease the weighted sample size and overlap between weighted samples. It can thus be seen a smoothing parameter, with higher values reducing the effect of fluctuations due to sampling error, but at the cost of reduced precision to detect meaningful differences.

In summary, LSEM is a sequential estimation of several SEMs based on a moving sample weight window. It allow researchers to fit models across age as a continuous moderator variable instead of using artificially created age groups, which are commonly used in cross-sectional personality development studies. Instead, we could fully make use of the information provided by a chronological age metric.

## *Method artifacts in examining differential and correlated change*

Findings based on self-reported personality levels may be affected by participants' response tendencies, such as acquiescent or socially desirable responding (Paulhus & Vazire, 2007). Acquiescent responding describes the tendency to agree with statements, independent of content. Strong acquiescence effects can potentially increase correlations between otherwise uncorrelated scales by contributing common response variance across unbalanced scales. Similarly, social desirable responding or self-evaluation tendencies can increase correlations between highly desirable (e.g. *intelligent* and *nice*) or undesirable items (e.g. *egoistical* and *irritable*) items—and decrease them between items of opposing desirability (e.g. *intelligent* and *egoistical*; Leising et al., 2015, 2020). However, desirability and item keying are heavily intertwined in the context of personality measurement (e.g. positive Conscientiousness, positive Agreeableness, and negative Neuroticism items being among the most desirable) and difficult to control for without removing relevant trait variance.

Conceptually, systematic response styles will potentially increase correlations between items or scales with similar properties (i.e. direction of keying, direction of desirability), thus modifying observed cross-sectional correlations. If the stability of the bias differs from the stability of the measured trait, the rank-order stability estimates will be affected. For instance, if acquiescence is less stable than the trait measured, the rank-order stability of the trait will be underestimated. Acquiescent responding seems to be relatively stable across time, with rank-order consistencies ranging from $r = .56$ (Billiet & Davidov, 2008), $r = .66$ (Danner et al., 2015), to $r = .77$ (Weijters et al., 2010) across time intervals of at least one year—thus being comparable to the reported personality rank-order stabilities. In combination with the small reported variance contribution of acquiescence (e.g. 5% of scale score variance; Danner et al., 2015), the effect on longitudinal change studies should be relatively weak. Unfortunately, we know very little about the stability of self-evaluation tendencies or socially desirable responding across such long time intervals. A recent study showed that self-evaluation tendencies can be attributed to differences in self-esteem (Leising et al., 2020) and could thus potentially be as stable as personality traits (Trzesniewski et al., 2003). A study examining the rank-order stability of a social desirability scale found a retest correlation of around $r = .66$ across 15 months (Haberecht et al., 2015). But as the items used in such scales are very similar to personality items, in particular Agreeableness and Conscientiousness, it is unclear to what degree these correlations reflect stability in personality traits or in self-evaluation biases. In addition, the strength of self-evaluation effect on the validity of personality measures is disputed (Barrick & Mount, 1996; Borkenau & Ostendorf, 1992; Holden & Passey, 2010; Piedmont et al., 2000).

## *The present study*

The goal of the present study was to examine differential and correlated personality change across time and age. Specifically, we focused on the following four guiding research questions: (a) Is change correlated across the Big Five personality traits? (b) Is the pattern of correlated personality change equal to the cross-sectional initial level correlations? Finally, extending previous studies on differential and correlated change, we examine how the (c) length of the time interval and (d) age at the first measurement occasion moderate the rank-order stability and change correlations. This would allow to examine whether potential time and age effects are specific to correlated change or that they apply to differential change as well.

To answer our research questions, we used two panel studies with broad age ranges and several measurement occasions, up to 18 years apart. This design has allowed us to study both the effect of age and time on rank-order stability and correlated change. To ensure the comparability of our results across a broad age range, we have selected two panel studies using the same personality measurement, but with only partly overlapping age ranges. We then used latent change score models (LCSMs) to examine intraindividual change and interindividual differences in intraindividual change (Ferrer & McArdle, 2010; McArdle, 2009; McArdle & Hamagami, 2001). This method has allowed us to estimate both initial trait level correlations and correlated change (Hertzog & Nesselroade, 2003). Third, we tested for measurement invariance across both time and age to ensure that the findings were not confounded by differences in the measurement process across these two moderators (Horn & McArdle, 1992; Meredith & Horn, 2001; Nye et al., 2016). Finally, we used LSEM (Briley et al., 2015; Hildebrandt et al., 2009, 2016; Hülür et al., 2011; Olaru et al., 2019) to consider whether and how rank-order stability and correlated personality change differs as a function of age.

## Methods

The analyses were exploratory and not preregistered. In this study, we reanalyzed publicly available and anonymous data from two panel studies. The analyses scripts and supplementary materials are available in an OSF repository (https://osf.io/x9u2j/).

### *Participants*

We used data from the Midlife in the United States (MIDUS) study (Barry, 2014), which is a longitudinal study of health and well-being in the United States with three measurement waves over a total of 19 years. In the first measurement wave in 1995, participants were recruited via telephone using a Random Digit Dial technique. Randomly selected households were contacted via telephone and within these households; one randomly selected family member was selected for participation (given the age restrictions). The second measurement wave was conducted roughly nine years later in 2004/2005 and the third one in 2013/2014. Data were collected via phone interviews and self-report questionnaires. For our analysis, we used respondents that participated in the personality assessment on least one measurement occasion. This criteria resulted in 3250 participants (1792 females) with an average age of 45.66 years ($SD = 11.40$; range = 20–74) on the first measurement occasion, 54.59 years of age ($SD = 11.35$; range = 30–84) on the second measurement occasion, and 63.69 years ($SD = 11.35$; range = 39–93) on the third measurement occasion.

In addition, we used data from the Health and Retirement Study (HRS; Juster & Suzman, 1995). The HRS is a longitudinal panel study of households in the United States with at least one member of 50 years of age or more. Participants in the study were members of the selected households (one or more member per household). The HRS started in 1992, and data were collected every two years since. New cohorts were added every six years to account for panel attrition. Personality was measured starting in 2006, and repeated measures were administered every four years. For this study, we used the data from the personality assessments in 2006, 2010, 2014, and 2018. We only used the data of persons who participated in at least one personality assessment ($N = 4145$; female = 2549). On the first measurement occasion, participants were on average 63.30 years old ($SD = 8.75$; range = 30–90), on the second measurement occasion 67.46 years old ($SD = 8.76$; range = 34–95), on the third measurement occasion 71.18 years old ($SD = 8.71$; range = 37–98), and on the fourth measurement 75.30 years old ($SD = 8.74$; range = 41–102).

### *Measures*

In both studies, personality was measured using the Midlife Development Inventory (MIDI) scales (Lachman & Weaver, 1997; Zimprich et al., 2012). This allowed us to compare the findings across the only partly overlapping age range of the two studies. The MIDI measures the Big Five personality factors with 25 adjective ratings (four to seven adjectives per trait). Each item was administered using a 4-point Likert scale ranging from *a lot (1)* to *not at all (4)*. For subsequent analysis, we have reversed the scales such that higher scores reflect a higher level on the trait. The adjectives used were as follows:

Neuroticism: *moody*, *worrying*, *nervous*, and *calm* (reverse keyed).

Extraversion: *outgoing*, *lively*, *active*, *talkative*, and *friendly*.[1]

Openness: *creative*, *imaginative*, *intelligent*, *curious*, *broadminded*, *sophisticated*, and *adventurous*.

Agreeableness: *helpful*, *warm*, *caring*, *softhearted*, and *sympathetic*.

Conscientiousness: *organized*, *responsible*, *hard-working*, and *careless* (reverse keyed).

### *Statistical analysis*

The analysis approach used in this study was as follows. We first established longitudinal latent variable models for each trait and tested for measurement invariance across the measurement occasions. We examined the rank-order stability of the constructs of interest based on the retest correlations estimated with correlated factor models. To examine correlated change over time, we established LCSMs for each personality trait (i.e. five univariate change score models) and each combination of personality factors (i.e. 10 dual change score models). Finally, we tested the potential moderating role of age on the rank-order stability and correlated change using LSEM (Hildebrandt et al., 2009, 2016). We provide a more detailed description in the following.

### *Correlated factor models*

To estimate differential stability, we specified longitudinal correlated factor models for each trait. Each factor was measured by the corresponding four to seven adjective ratings collected on the measurement occasion. To scale the factors, the first factor loading was constrained to 1 and factor means to 0. Residuals of the same item were allowed to correlate across time (Little, 2013). For the Openness model, we included residual correlations within the measurement occasions between *sophisticated* and *intelligent*, as well as *creative* and *imaginative*, because the correlation between these items could not be fully explained by the common factor (see also Zimprich et al., 2012). We also included a residual correlation between *active* and *lively* in the Extraversion model for the same reason.

### Latent baseline change model

To differentiate between initial personality levels and subsequent change as a function of time and age, we modified the correlated factor model by adding a regression of the second, third, and fourth (only in HRS) measurement occasion factors on the first measurement occasion factor with a weight of $\beta = 1$. The residuals of the regressions represent the personality change across 4, 8, and 12 (HRS), or 9 and 18 years (MIDUS). This model is similar to LCSM (McArdle, 2009), but all measurement occasion factors are compared to the first measurement occasion instead of the previous measurement occasion factor. We chose this type of model, because it allowed us to achieve varying time intervals for personality change within one model. This model is also less restrictive concerning the shape of the change compared to latent growth curve models (McArdle, 2009), which would only allow us to model linear change trajectories across the three measurement occasions in the MIDUS dataset. To estimate correlated change between the personality traits, we created bivariate change score models by combining two change score models and additionally estimating the factor correlations (see Figure 1). This resulted in five univariate and ten bivariate change score models for each study (i.e. HRS and MIDUS). To examine whether the change score correlations were inflated due to regression to the mean, we compared the findings to a latent baseline change model in which the change scores were regressed on the first measurement occasion (OSF Figure 1B). We further compared it to a cross-lagged panel model (McArdle, 2009) with unconstrained regression paths, to check how the regression constraints and comparison to the baseline instead of previous measurement occasion affected the results (OSF Figure 1C).

### Model estimation and measurement invariance

Models were estimated in R using the lavaan Package (Rosseel, 2012) and *full information maximum likelihood* (FIML) estimation to account for missing data. To ensure that the choice of estimation did not affect the results, we also estimated models with the *weighted least squares means and variance adjusted* (WLSMV) estimator. The relevant parameters for this study were equivalent across both estimation methods. We thus used FIML estimation because it is more efficient at estimating the large number of models used in this study and is appropriate for

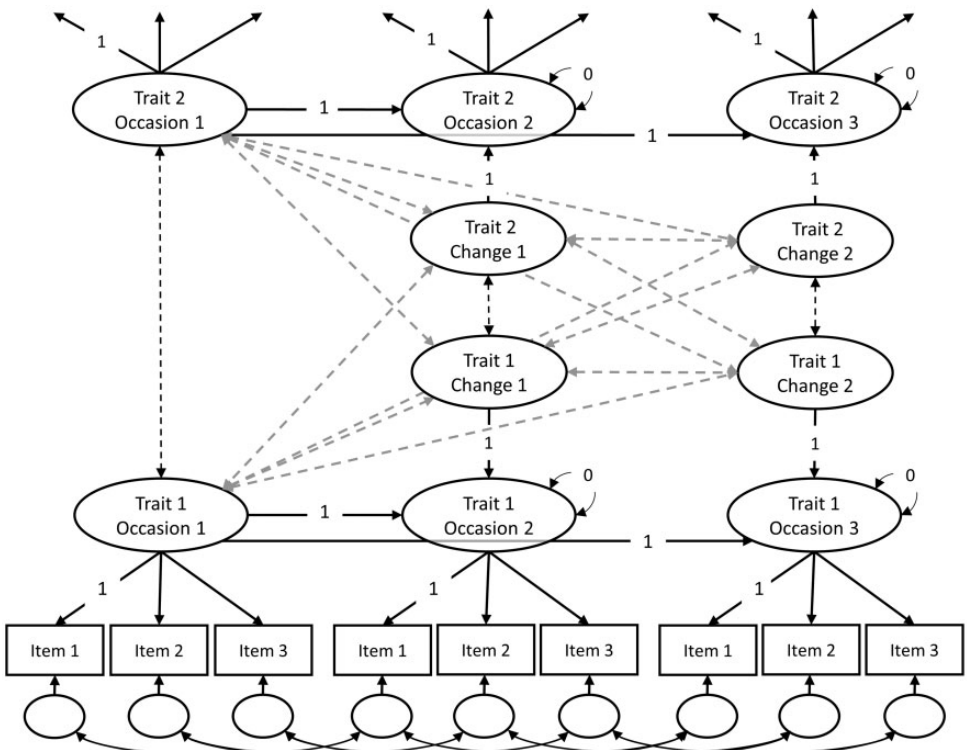


**Figure 1.** Latent change score model. This is a simplified illustration of the bivariate change score models used in this study, which ranged from four to seven indicators per measurement occasion factor and included a fourth measurement occasion factor in the HRS sample. Dashed double-arrowed lines represent estimated factor correlations. The black dashed lines are the parameters of interest in this study. We did not estimate a factor mean structure (i.e. factor means were constrained to 0).

indicators with at least four levels (Beauducel & Herzberg, 2006; Rhemtulla et al., 2012).

We evaluated overall model fit with a combination of the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR) based on common standards (acceptable/good fit: CFI $\geq$ .90/.95; RMSEA$\leq$.08/.06; SRMR$\leq$.08/.06; Bentler, 1990; Hu & Bentler, 1999). We tested for the measurement invariance of the model across measurement occasions and participants' age (Little, 2013; Meredith, 1993). More specifically, we compared model fit of (a) a model without parameter constraints (i.e. configural invariance), to (b) a model with factor loadings constrained to equality across time and age (i.e. metric invariance), to (c) a model with additionally constrained item intercepts across time and age (i.e. scalar measurement invariance). As we were only interested in factor correlations, metric measurement invariance across measurement occasions and age samples was sufficient (Little, 2013). Measurement invariance was tested by evaluating the increase in model misfit between nested models based on a cutoff of $\Delta$CFI = –.010, $\Delta$RMSEA = .015 and $\Delta$SRMR = .030/.015 (metric/scalar measurement invariance; Chen, 2007; Cheung & Rensvold, 2002). If a change in fit was greater than these cutoffs, then measurement invariance over time was not supported. The equivalence of single model parameters over time (e.g. factor correlations between change factors across time) was tested by specifying difference parameters in the model and testing whether these parameters differed significantly from zero.

## Local structural equation modeling

In the MIDUS dataset, we estimated the models ranging from 25 to 65 years. In the HRS data, we estimated models from 50 to 80 years of age. Participants outside of these age ranges are still included in the model estimation because of the Gaussian weighing kernel applied by LSEM. We chose these age ranges because the resulting weighted sample size was too small for robust parameter estimates outside this range. In line with the recommendations in the literature (Hildebrandt et al., 2016), we used a bandwidth value of $h = 2$, which provides a good trade-off between resulting sample size and power to detect parameter differences across the moderator. LSEM was run with the *lsem.estimate*-function in the R package *sirt* (Robitzsch, 2019). We estimated the models in five year increments because of a high computation load of smaller steps in the permutation test described below.

To test for measurement invariance across age in LSEM, we estimated the longitudinally measurement invariant model without and with additional constraints across age in LSEM. To derive global fit indices in LSEM that can be used for measurement invariance testing, we used the joint estimation procedure implemented in the *lsem.estimate*-function (Robitzsch, 2019). This approach uses a procedure similar to MGCFA to estimate the model simultaneously across the weighted samples (instead of separately in the default LSEM application). In the default settings, LSEM estimates the models sequentially across the age points. In the joint estimation, each weighted age sample is treated like an independent group in MGCFA and a common likelihood function is maximized. This approach can be used to constrain and estimate parameters across the age points (e.g. for measurement invariance testing). Whereas the $\chi^2$ statistic and standard errors for the estimates are not interpretable due to the increased overall sample size, goodness-of-fit indices and parameter estimates are not affected by treating the weighted age samples as independent. We thus used aforementioned model fit cutoffs to test for measurement invariance across age (Chen, 2007; Cheung & Rensvold, 2002).

To test the moderation effect of age on single model parameters of the LCSMs, we used a permutation test (Hildebrandt et al., 2016; Hülür et al., 2011; Schroeders et al., 2015). The permutation test resembles traditional significance testing approaches, in which the parameter values are tested against a distribution that can be expected to occur because of sampling error. To create such a distribution, the permutation test creates 1000 resampled copies of the dataset (with default settings). Within each dataset, the moderator values are randomly shuffled across individuals (Hülür et al., 2011; Jorgensen et al., 2018). By doing so, all systematic moderation effects are removed from the data. LSEM is then run on each dataset to derive the model parameters. This procedure results in a distribution of estimates for each parameter in which the estimates are independent of the moderator. The original LSEM parameter estimates are then compared to the corresponding distribution under the null hypothesis, which allows for the identification of significant moderation effects on each parameter.

## Method artifacts in studies on differential and correlated change

To better understand how systematic variations in the responses across time affect the parameters examined in this study, and also how these are interconnected (e.g. does a higher cross-sectional correlation result in higher change correlations?), we conducted a small simulation study (for a detailed description and results, see https://osf.io/x9u2j/). More specifically, we simulated two scales scores at two measurement occasions. Each scale score consisted of a trait component that was only correlated across time (but uncorrelated within time points) and a shared variance (SV) component that was equivalent within a measurement occasion, but allowed to vary across time. This SV can conceptually represent the common effect of mechanisms on two traits (e.g. a normative life event affecting both traits simultaneously) but also content-independent response styles. We then investigated how the (a) test–retest, (b) cross-sectional, (c) cross-lagged, and (d) change score correlations varied as a function of (1) the rank-order stability of the *trait variance* ranging from $r = .50$ to $r = .90$, (2) the rank-order stability of the added SV ranging from $r = .00$ to $r = 1.00$, and (3) the proportion of the SV ranging from 0% to 50% of the total scale score variance. In the following, we summarize the main findings (for detailed results see https://osf.io/x9u2j/).

Not surprisingly, test–retest correlations (i.e. differential stability) depended on the stability of both variance components and their relative contribution to the overall scores. Assuming that response styles are about as (or only slightly less) stable as the personality traits, the differential stability estimates should only be weakly biased. Because of the previously reported similarities between cross-sectional and change correlations, we were interested in whether cross-sectional correlations necessarily imply correlated change. Change score correlations were positively related to cross-sectional correlations, because both parameters depend on the strength of the SV at one or more time points. As such, correlated change can only be found if the scores correlate at least at one time point. However, it is possible to find no correlated change despite cross-sectional correlations at both time points (i.e. if only the unshared variance changes across time). Change score correlations also increased with increasing variability of the SV across time and increasing stability of the *trait variance*. We were able to replicate the correlations reported in previous studies using a combination of a relatively stable SV (i.e. test–retest correlation from .60 to .80), and a compensatory effect of bias strength (40% to 50% of variance contribution) for variable traits (test–retest correlation of .50 to .60) or lower bias strength (20% to 30% of variance explained) for more stable traits (test–retest correlation of .70 to .90). While this example shows that it is possible to find cross-sectional correlations and correlated change by adding a SV component with moderate to high stability to otherwise uncorrelated scores, the nature of this component is unclear. It most likely represents a combination of the overlap between the measured traits, bidirectional feedback loops, external occasion specific causes that affect both traits simultaneously, and the aforementioned response biases.

## Results

### *Model fit, measurement invariance, and reliability*

We first tested the longitudinal models for measurement invariance across time and age (LSEM) to ensure that model parameters were comparable. The univariate models yielded an acceptable absolute model fit and metric measurement invariance across both time and age (see Table 1). Based on the common cutoff criteria, scalar measurement invariance was not given for Neuroticism ($\Delta CFI = -.011$), Extraversion ($\Delta CFI = -.012$), and Conscientiousness ($\Delta CFI = -.025$) in the MIDUS sample. Conscientiousness was the only factor affected by a lack of scalar measurement invariance in the HRS sample ($\Delta CFI = -.029$). The lack of item intercept invariance in the MIDUS sample can be attributed to the longer time intervals and broader age span in this sample. However, the MIDI Conscientiousness scale seems to be problematic for mean-level comparisons across measurement occasions and age in general. For the current analysis of correlational patterns, metric measurement invariance is sufficient (Little, 2013), as we were interested in interindividual differences in change, but not the mean-levels or precise change scores. Metric invariance was also achieved for all bivariate change score models (see OSF Table 1; https://osf.io/x9u2j/). Correlated factor models and the LCSMs were equivalent regarding model fit and degrees of freedom. All scales except for Conscientiousness—which was measured with only four items—yielded acceptable internal consistencies (i.e. Cronbach's $\alpha$/McDonald's $\omega > .70$; see Table 1).

### *Differential change*

We then examined differential change as a function of varying time intervals and participants' age. The rank-order stability of the traits—estimated in correlated factor models on the full samples—is presented in Table 2. Within the MIDUS study, we were able to estimate two 9-year test–retest correlations, and one 18-year test–retest correlation for all personality factors. In the HRS study, three 4-year correlations, two 8-year correlations, and one 12-year test–retest correlation were estimated for each personality factor. In both studies, the test–retest correlations for each trait were statistically equivalent across the comparable time intervals (e.g. the three 4-year test–retest correlations for Agreeableness were equivalent in the HRS study). With increasing time intervals, the rank-order stability decreased significantly for all traits ($ps < .01$) but was relatively high even after 12 years (HRS: $r = .64$ to .76) and 18 years (MIDUS: $r = .65$ to .80), respectively. Overall, the rank-order stabilities were largely in the range of the stability estimates reported in the meta-analytic reviews (Ferguson, 2010; Roberts

**Table 1.** Univariate model fit, reliability, and measurement invariance across time and age.

| | α/ω | MI level | *df* | $\chi^2$ value | Across time CFI | Across time RMSEA | Across time SRMR | Across age (LSEM) CFI | Across age (LSEM) RMSEA | Across age (LSEM) SRMR |
|---|---|---|---|---|---|---|---|---|---|---|
| MIDUS | | | | | | | | | | |
| N | .74/.75 | Configural | 39 | 134 | .993 | .027 | .029 | .997 | .018 | .034 |
| | | Metric | 45 | 148 | .993 | .027 | .032 | .997 | .017 | .036 |
| | | Scalar | 51 | 308 | .982 | .039 | .041 | .976 | .045 | .053 |
| E | .73/.72 | Configural | 36 | 424 | .977 | .058 | .043 | .977 | .056 | .050 |
| | | Metric | 42 | 437 | .976 | .054 | .044 | .977 | .055 | .052 |
| | | Scalar | 48 | 648 | .964 | .062 | .050 | .963 | .063 | .060 |
| O | .77/.75 | Configural | 159 | 1809 | .941 | .057 | .056 | .943 | .057 | .063 |
| | | Metric | 171 | 1816 | .941 | .054 | .057 | .943 | .056 | .066 |
| | | Scalar | 183 | 1872 | .940 | .053 | .051 | .939 | .056 | .068 |
| A | .79/.80 | Configural | 72 | 430 | .981 | .039 | .027 | .985 | .034 | .032 |
| | | Metric | 80 | 447 | .980 | .038 | .029 | .984 | .035 | .038 |
| | | Scalar | 88 | 533 | .976 | .039 | .030 | .977 | .039 | .041 |
| C | .57/.56 | Configural | 39 | 169 | .986 | .032 | .030 | .985 | .033 | .039 |
| | | Metric | 45 | 210 | .983 | .034 | .033 | .985 | .032 | .041 |
| | | Scalar | 51 | 451 | .958 | .049 | .045 | .958 | .049 | .050 |
| HRS | | | | | | | | | | |
| N | .71/.72 | Configural | 74 | 328 | .989 | .029 | .037 | .992 | .024 | .042 |
| | | Metric | 83 | 342 | .988 | .027 | .039 | .993 | .024 | .043 |
| | | Scalar | 92 | 410 | .986 | .029 | .040 | .985 | .031 | .049 |
| E | .72/.71 | Configural | 70 | 719 | .978 | .047 | .050 | .978 | .048 | .055 |
| | | Metric | 79 | 732 | .977 | .045 | .050 | .978 | .047 | .056 |
| | | Scalar | 88 | 891 | .972 | .047 | .053 | .973 | .050 | .059 |
| O | .80/.78 | Configural | 294 | 2075 | .964 | .038 | .040 | .967 | .039 | .047 |
| | | Metric | 312 | 2110 | .964 | .037 | .041 | .967 | .039 | .048 |
| | | Scalar | 330 | 2296 | .960 | .038 | .042 | .964 | .039 | .050 |
| A | .79/.79 | Configural | 134 | 512 | .987 | .026 | .024 | .986 | .028 | .034 |
| | | Metric | 146 | 547 | .986 | .026 | .026 | .986 | .028 | .036 |
| | | Scalar | 158 | 676 | .982 | .028 | .027 | .979 | .033 | .039 |
| C | .56/.57 | Configural | 74 | 248 | .990 | .024 | .028 | .990 | .024 | .040 |
| | | Metric | 83 | 295 | .987 | .025 | .033 | .990 | .024 | .041 |
| | | Scalar | 92 | 787 | .958 | .043 | .045 | .956 | .047 | .055 |

α/ω: Cronbach's internal consistency alpha/McDonald's factor saturation omega; across time: model fit for the longitudinal model estimated on the full sample; across age (LSEM): model fit for the longitudinal model estimated across age using joint estimation in local structural equation modeling; MI level: measurement invariance level; configural: no additional constraints across time or age; metric: factor loading equality constraints across time and age (in LSEM); scalar: factor loading and item intercept equality constraints across time and age (in LSEM); *df*: degrees of freedom; CFI: comparative fit index; RMSEA [90% CI]: root mean square error of approximation with 90% confidence interval; SRMR: standardized root mean square residual; N: Neuroticism; E: Extraversion; O: Openness for Experiences; A: Agreeableness; C: Conscientiousness. One and two residual correlations within measurement occasions were included in the Extraversion and Openness model, respectively (see Methods section). We first tested longitudinal measurement invariance in the full sample (across time). To test for measurement invariance across age, we estimated the longitudinally measurement invariant model with and without additional equality constraints across age in LSEM. Degrees of freedom and $\chi^2$ values for LSEM are inflated due to treating the weighted samples as independent and are thus not reported.

& DelVecchio, 2000). Test–retest correlations across comparable time intervals were slightly higher in the MIDUS sample (MIDUS: average $r_{\text{test–retest9}} = .78$; HRS: average $r_{\text{test–retest8}} = .74$).

We used LSEM to examine the moderation effect of age on the test–retest correlation of the traits (see Figure 2). After correcting for the influence of age, the average test–retest correlations were equivalent to the full model and age had no significant moderation effect on the rank-order stability (Table 2). The effect of age on the test–retest correlations was thus less pronounced than differences across time intervals, and we found no systematic support for the *cumulative stability principle* or for *diversification* in old age.

### Correlations between initial trait levels

Next, we examined correlated change in the dual change score models and compared these to the baseline trait correlations. Because of the large number of correlations, we will first discuss the findings for the correlations at the first measurement occasion (i.e. cross-sectional correlations) and present correlated change patterns in the following section. Factor correlations are presented in Table 3. Across both studies, Neuroticism was uncorrelated with Agreeableness (MIDUS $r = .04$; HRS $r = -.03$) and yielded only small negative correlations with Extraversion, Openness, and Conscientiousness (MIDUS average

**Table 2.** Differential change across varying time intervals.

| | | MIDUS | | HRS | | |
|---|---|---|---|---|---|---|
| | | 9 years (T1–T2/T2–T3) | 18 years | 4 years (T1–T2/T2–T3/T3–T4) | 8 years (T1–T3/T2–T4) | 12 years |
| Neuroticism | r | .74 (.73/.74) | .65 | .75 (.74/.76/.75) | .69 (.66/.71) | .64 |
| | *p* | *.399 (.029/.984)* | *.522* | *.202 (.896/.257/.266)* | *.107 (.059/.566)* | *.184* |
| Extraversion | r | .84 (.83/.85) | .80 | .84 (.82/.85/.85) | .79 (.77/.81) | .75 |
| | *p* | *.589 (.331/.729)* | *.060* | *.494 (.707/.307/.772)* | *.970 (.743/.901)* | *.634* |
| Openness | r | .80 (.80/.80) | .74 | .83 (.80/.84/.84) | .79 (.79/.79) | .76 |
| | *p* | *.735 (.289/.582)* | *.032* | *.956 (.916/.548/.459)* | *.187 (.352/.306)* | *.076* |
| Agreeableness | r | .74 (.73/.76) | .70 | .75 (.75/.76/.75) | .72 (.73/.70) | .69 |
| | *p* | *.190 (.190/.383)* | *.400* | *.196 (.104/.842/.214)* | *.174 (.563/.140)* | *.302* |
| Conscientiousness | r | .76 (.77/.74). | .70 | .78 (.78/.77/.79) | .73 (.74/.71) | .66 |
| | *p* | *896 (.180/.548)* | *.856* | *.149 (.288/.713/.227)* | *.107 (.074/.616)* | *.281* |

*r*: test–retest correlations; *p*: *p* values for the age moderation in local structural equation modeling. Values outside the parentheses represent the average test–retest correlations across time intervals. Within parentheses, the pairwise test–retest correlations are presented. Numbers in *italics* represent the *p* values of the age moderation with local structural equation modeling.

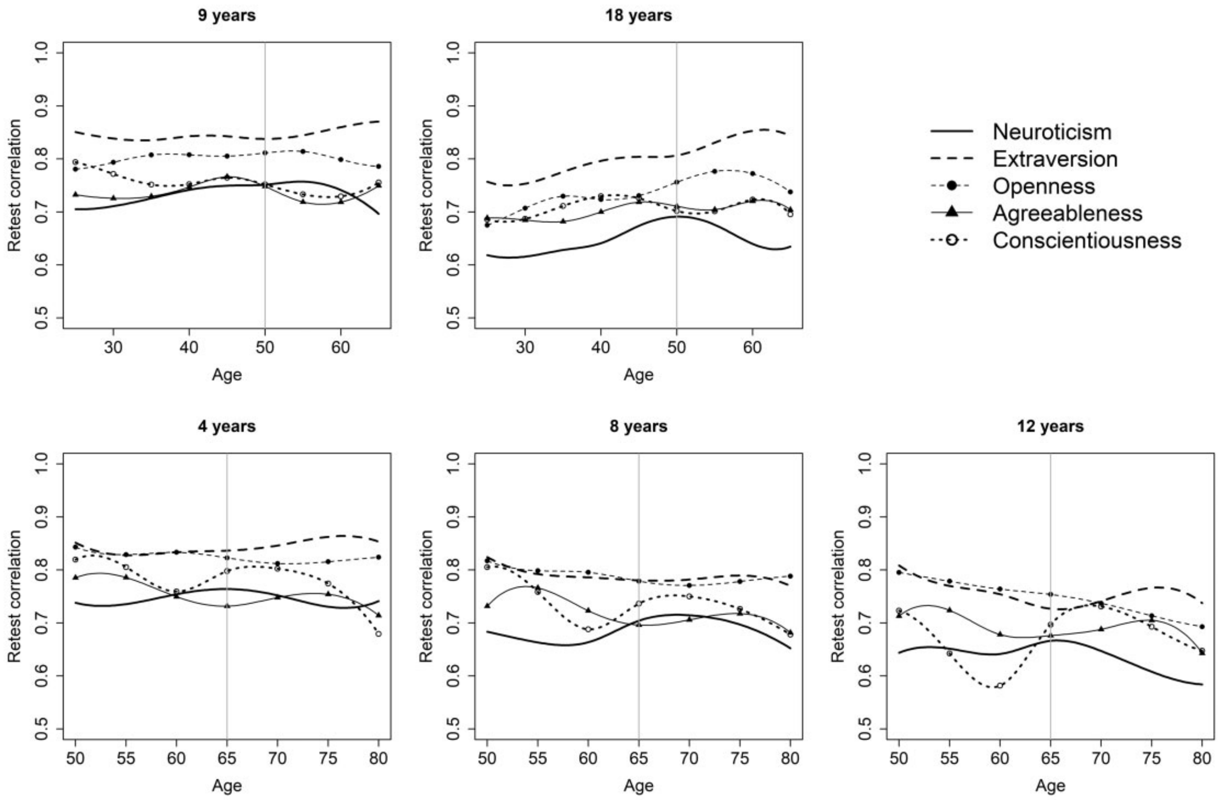


**Figure 2.** Differential change across age. The five horizontal lines per plot show the average test–retest correlations between factors estimated with local structural equation modeling across age. The upper panels represent the test–retest correlations found in the MIDUS sample. The lower panels show the HRS test–retest correlations. Gray vertical lines indicate the age overlap between studies.

$r = -.15$; HRS average $r = -.19$; $ps < .001$). In contrast, Extraversion, Openness, Agreeableness, and Conscientiousness showed a medium to large positive factor correlation with each other at the first measurement occasion (MIDUS average $r_{EOAC} = .47$; HRS average $r_{EOAC} = .61$; $ps < .001$).

Using LSEM, we examined the moderation effect of chronological age on the initial measurement factor correlations (for exact correlations and *p* values, see OSF Table 2: https://osf.io/x9u2j/). The left panels in Figure 3 show the correlations between Neuroticism and the other personality factors over age in both studies. Generally, correlations for Neuroticism remained stable (i.e. small and negative or zero). We found an increase in the Neuroticism-Agreeableness correlation from –.12 to .22 in the HRS sample ($p = .038$), but these correlations did not differ significantly from zero at any age. This

**Table 3.** Level and change score correlations.

| MIDUS | Level | 9 years | 18 years | HRS | Level | 4 years | 8 years | 12 years |
|---|---|---|---|---|---|---|---|---|
| Model with change-level correlation | | | | | | | | |
| N–E | –.14 | **–.18** | **–.13** | N–E | **–.21** | –.10 | **–.18** | **–.14** |
| N–O | **–.15** | **–.15** | **–.14** | N–O | **–.18** | –.09 | **–.11** | **–.14** |
| N–A | .04 | –.04 | –.07 | N–A | –.03 | .05 | .00 | –.05 |
| N–C | **–.12** | –.12 | **–.16** | N–C | **–.17** | –.09 | –.08 | –.09 |
| E–O | **.62** | **.81** | **.73** | E–O | **.68** | **.88** | **.89** | **.88** |
| E–A | **.56** | **.68** | **.63** | E–A | **.66** | **.81** | **.72** | **.78** |
| E–C | **.35** | **.46** | **.46** | E–C | **.59** | **.74** | **.70** | **.78** |
| O–A | **.41** | **.59** | **.60** | O–A | **.48** | **.69** | **.68** | **.72** |
| O–C | **.40** | **.53** | **.61** | O–C | **.58** | **.77** | **.76** | **.80** |
| A–C | **.46** | **.53** | **.58** | A–C | **.63** | **.87** | **.83** | **.84** |
| Model with change-level regression | | | | | | | | |
| N–E | **–.14** | **–.19** | **–.13** | N–E | **–.21** | **–.14** | **–.26** | **–.20** |
| N–O | **–.15** | **–.20** | **–.15** | N–O | **–.18** | **–.12** | **–.15** | **–.17** |
| N–A | .04 | –.09 | –.08 | N–A | –.03 | .03 | –.05 | –.07 |
| N–C | **–.12** | **–.15** | **–.20** | N–C | **–.17** | **–.14** | **–.16** | **–.17** |
| E–O | **.62** | **.80** | **.73** | E–O | **.68** | **.87** | **.89** | **.89** |
| E–A | **.56** | **.67** | **.63** | E–A | **.66** | **.81** | **.73** | **.78** |
| E–C | **.35** | **.43** | **.45** | E–C | **.59** | **.73** | **.71** | **.78** |
| O–A | **.41** | **.56** | **.60** | O–A | **.48** | **.68** | **.69** | **.73** |
| O–C | **.40** | **.51** | **.60** | O–C | **.58** | **.77** | **.77** | **.81** |
| A–C | **.46** | **.50** | **.57** | A–C | **.63** | **.86** | **.83** | **.83** |

MIDUS: Midlife in the United States; HRS: Health and Retirement Study; N: Neuroticism; E: Extraversion; O: Openness; A: Agreeableness; C: Conscientiousness; Level: correlation between first measurement occasion factors; X years: correlations between the X-year change scores. The correlations were estimated in bivariate baseline change models with correlations/regressions between the first measurement occasion (level) and the change scores. Parameters marked in bold are significant at the $p < .001$ level.

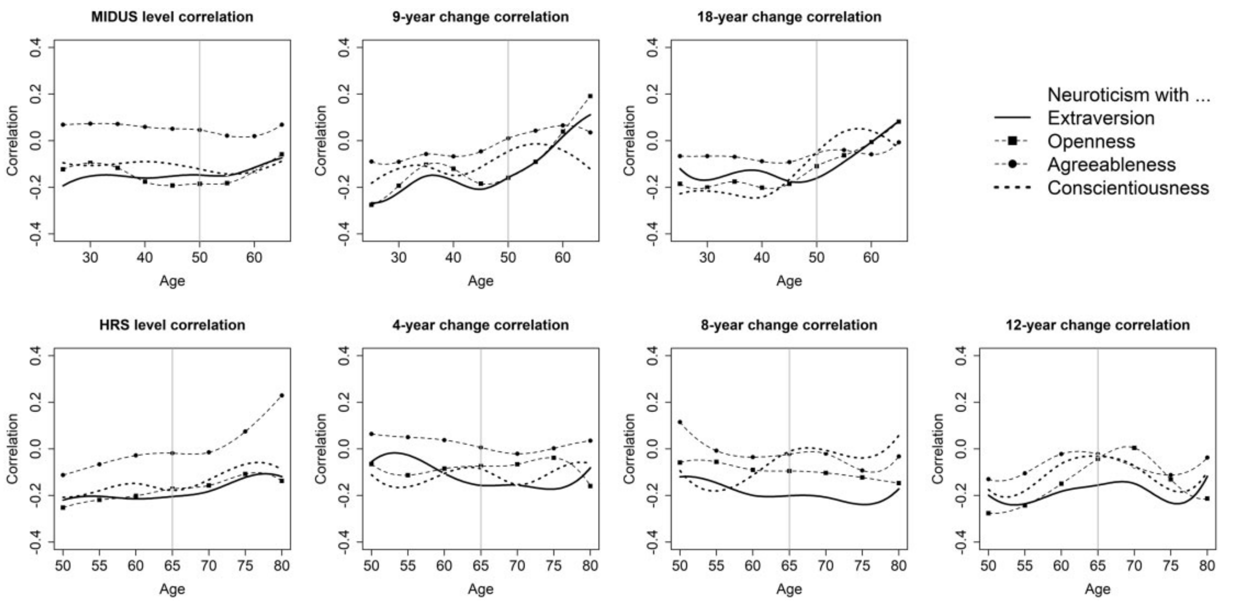


**Figure 3.** Correlated neuroticism change across age. The four horizontal lines per plot show the correlations between corresponding change factors estimated with local structural equation modeling across age at the first measurement occasion. The upper panels represent the correlations found in the MIDUS sample. The lower panels show the HRS correlations. Gray vertical lines indicate the age overlap between studies.
MIDUS: Midlife in the United States; HRS: Health and Retirement Study.

finding suggests that the effect—despite seeming descriptively large—is not significant and Neuroticism and Agreeableness remained uncorrelated. The leftmost panels of Figure 4 show the correlations between the Extraversion, Openness, Agreeableness, and Conscientiousness baseline levels as a function of age ($p_{EA} = .004$; $p_{EC} = .018$; $p_{OC} = .009$; $p_{AC} = .009$). For these four factors, we found an increase in the correlations in the MIDUS sample, in particular from 25 to 50 years of age. Initial correlations increased from an average $r = .37$ at age 25 to an average $r = .50$ at age 50 and remained relatively stable afterward (age 65 average $r = .54$). This effect was strongest for the Conscientiousness correlations, which increased by $\Delta r = .22$ from age 25 to 65 (compared to an average $\Delta r = .12$ for the other traits).

## Correlated change

We examined the correlations between change factors and compared them to the baseline trait correlations (see Table 3; for the full correlation table between all factors, see OSF Tables 4 and 5). To examine whether the change score correlations were inflated due to regression to the mean, we compared a change score model with a correlation between level and change to a model with a regression of all change scores onto the levels of both traits (see Table 3; for the cross-lagged panel results, see OSF Table 3). The change score correlations were comparable across all three models, suggesting that the change score correlations were not affected by a regression to the mean. We will thus discuss the findings based on the model with level-change correlation in the following. The change factor correlations between Neuroticism and the other traits were stable over time and statistically equivalent to the correlations at the first measurement occasion. The change factor correlations of the other four personality were also similar to the initial trait correlations, albeit on average slightly higher by $\Delta r = .13$ (MIDUS; average change $r = .60$) and $\Delta r = .19$ (HRS; average change $r = .79$). Over time, we did not find any differences between the change factor correlations. The change factor intercorrelations thus seem to be stable, even over long time intervals. This finding is particularly interesting given the decreasing rank-order stability over time (Table 2). In general, correlated change between Openness, Conscientiousness, Extraversion, and Agreeableness was very high. For example, a common change factor was able to explain 61% (MIDUS nine-year change) and 76% (HRS eight-year change) of the change score variance.

Using LSEM, we examined the moderation effect of chronological age on the change factor correlations. Figure 3 shows the correlations between Neuroticism change and the other change factors across age in both studies. Similar to the baseline level correlations, change correlations for Neuroticism remained stable across age. The only exception was the 4-year change correlation with Extraversion ($p = .007$) and 18-year change correlation with Conscientiousness ($p = .006$). Figure 4 shows the correlations between the Extraversion, Openness, Agreeableness, and Conscientiousness change factors as a function of age. For these four factors, we found an increase in correlated change in the MIDUS sample ($p_{\text{EC-9}} = .016$; $p_{\text{OC-9}} = .008$; $p_{\text{AC-9}} = .002$). Specifically, we found an increase in the average 9-year change factor correlations between the four traits from $r = .41$ to $r = .76$ across 40 years of age. Again, this effect was most pronounced for Conscientiousness, with nine-year change

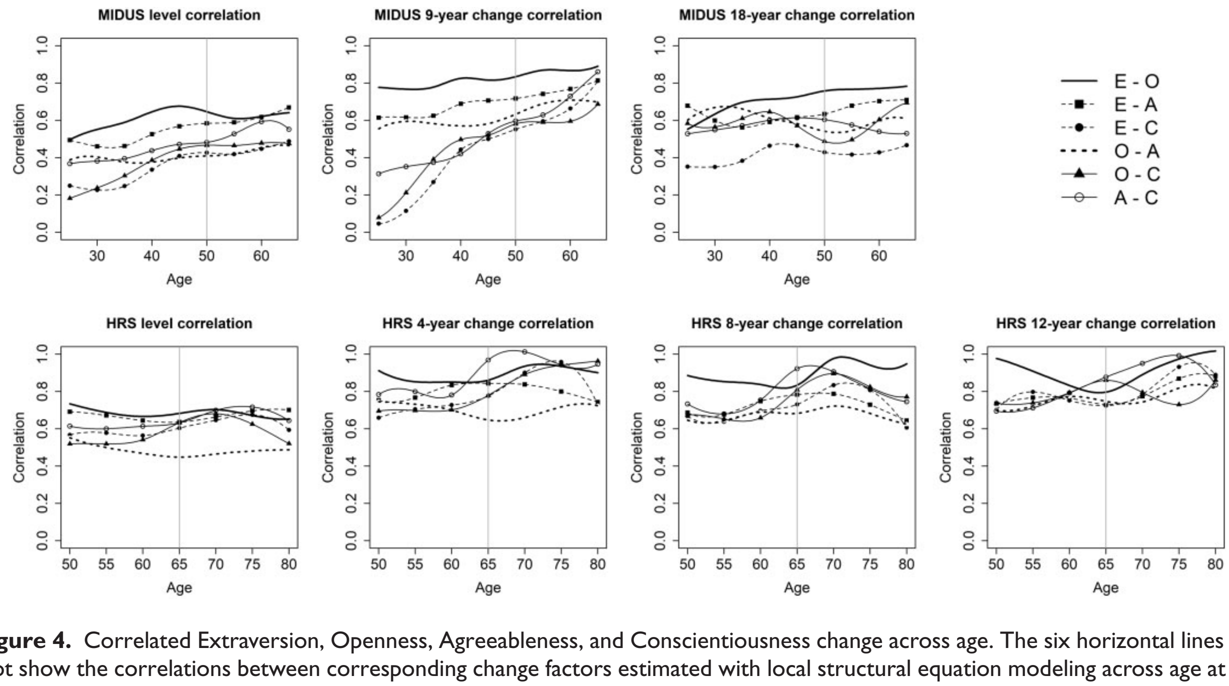


**Figure 4.** Correlated Extraversion, Openness, Agreeableness, and Conscientiousness change across age. The six horizontal lines per plot show the correlations between corresponding change factors estimated with local structural equation modeling across age at the first measurement occasion. The upper panels represent the correlations found in the MIDUS sample. The lower panels show the HRS correlations. Gray vertical lines indicate the age overlap between studies.
MIDUS: Midlife in the United States; HRS: Health and Retirement Study.

correlations increasing by $\Delta r = .54$. In the HRS study, the change correlations increased even further ($p_{EA\text{-}4} = .029$; $p_{EC\text{-}4/12} = .012/.043$; $p_{OC\text{-}8/12} = .005/<.001$; $p_{AC\text{-}4/12} = .024/.013$), nearly approaching equivalence in old age. As indicated by the age moderation, the higher trait and change factor correlations in the HRS study can be explained by the age differences between the studies (MIDUS mean age = 45.66 years, HRS mean age = 63.30 years).

In summary, these results indicate a strong age-dedifferentiation in the change processes across all Big Five traits but Neuroticism. This is particularly interesting given that the degree of differential change remained stable across this age span (Figure 2). As such, the strong correlated change does not seem to be the results of a small number of people changing systematically in old age. We also examined the change score variances across age (see OSF Figure 1) to ensure that the interindividual differences in intraindividual change were substantial across the entire age range (see also Schwaba & Bleidorn, 2017). We did find an age-associated decrease in the MIDUS change variances of Neuroticism (9/18 years *SD*: 0.40 to 0.27/0.42 to 0.32), Extraversion (9/18 years *SD*: 0.40 to 0.29/0.47 to 0.34), and Openness (9/18 years *SD*: 0.43 to 0.32/0.45 to 0.38). However, we found no systematic effect across age in the HRS sample. The overall change score variances were also similar between these two samples (see OSF Tables 3, 4 and OSF Figure 1), also suggesting that the higher change correlations in old age were not based on a small number of individuals that changed or relatively small change processes.

## Discussion

The goal of this study was to examine personality development through the lens of differential and correlated change. Whereas the perspective of differential change has received much attention in the field of personality development, correlated personality change is an underrepresented perspective of change and stability. The goal of this study was to demonstrate the importance and uniqueness of correlated change to investigate how basic personality processes unfold over time across the adult lifespan. Moreover, this work significantly expanded on previous research by comparing test–retest and change correlations across different time intervals and broad age ranges. Most notably, we made use of LSEM (Hildebrandt et al., 2009, 2016; Olaru et al., 2019) to maintain the continuous nature of age, rather than using artificially categorized age groups created for comparison. In the following, we will first discuss our main findings on differential change, followed by the cross-sectional correlations between initial trait levels and the findings on correlated change.

Consistent with previous research, we found relatively high levels of rank-order stability over time during adulthood as well as decreases in the test–retest correlations with increasing time between measurements (Anusic & Schimmack, 2016; Ferguson, 2010; Roberts & DelVecchio, 2000). Surprisingly, we found neither support for the cumulative stability principle with increases in differential stability (Caspi et al., 2005; Roberts & DelVecchio, 2000; Roberts & Wood, 2006) nor for an inverted U-shape of differential change with a diversification in old age (Lucas & Donellan, 2011; Specht et al., 2011; Wortman et al., 2012). However, the findings with respect to a decrease after the peak in the sixth decade of life were based on four-year test–retest correlations of two panel studies—the German Socio-Economic Panel Study (GSOEP) and The Household, Income, and Labor Dynamics in Australia (HILDA). A replication of these studies with LSEM only partly supported the decreasing rank-order stability in old age (Wagner et al., 2019). Unfortunately, only the HRS sample allowed us to estimate models above a mean age of 70 years, which was the range in which the decline could be observed (Lucas & Donellan, 2011; Wortman et al., 2012). Further analyses of differential change across age are required to evaluate the principle of *diversification* in old age.

We found strong support for correlated change in this study, particularly in late adulthood. All change factors except for Neuroticism were highly correlated with each other. One of the central findings of this study was that correlated change across all time intervals was very similar to the cross-sectional correlations at the first measurement occasion, providing evidence for “intercorrelations stationarity” (i.e. the change correlations converge toward the cross-sectional correlations; cf. Hofer et al., 2006) with respect to personality traits (Allemand & Martin, 2016). This finding was particularly evident when comparing the small negative Neuroticism correlations with the strong positive correlations between the other Big Five factors. This is also supported by previous studies on correlated personality change, which also reported similarities between correlated change and cross-sectional correlations (Allemand et al., 2007, 2008; Klimstra et al., 2013; Mõttus et al., 2012). To return to the question of whether correlated change is the product of narrow or broadly acting mechanisms, it seems to be affected by the same mechanisms that are responsible for the initially observed trait correlations. Because these correlations strongly depend on the personality model and measurement used, we will not discuss single correlations in detail but instead focus on the more general level of correlated change and “intercorrelations stationarity.”

One possible explanation for the similarities between cross-sectional and change correlations is that trait levels are in fact an accumulation of personality change up to the current moment. The youngest participants in this study were 20 years old and had thus already experienced 20 years of (correlated) personality change. This change is the product of, among others, biological maturation, genetic factors, social roles, culture, life events, health issues, social environments, and person–environment interactions (Caspi & Roberts, 2001; Fraley & Roberts, 2005; Kandler et al., 2015; Roberts & Helson, 1997; Roberts & Wood, 2006; Roberts et al., 2008; Wood & Denissen, 2015; Wrzus & Roberts, 2017). These broadly acting mechanisms affect several traits simultaneously, either directly at the trait level (e.g. McCrae et al., 2000) or by changing the trait-related behaviors, emotions, or thoughts (e.g. Wrzus & Roberts, 2017). For instance, changes in Behavioral Activation or Inhibition (Cloninger et al., 1991), or the dopaminergic system (DeYoung & Gray, 2009), might affect Extraversion and Openness in tandem, which in turn results in positive cross-sectional correlations between these traits. Such a cumulation of various correlated change processes across the lifespan might also explain the slightly increasing baseline level correlations we found in the MIDUS study.

Alternatively, it is also conceivable that change correlations are a product of some form of transfer effects across personality traits. This transfer may also work directly at the level of personality-related behaviors (Wrzus & Roberts, 2017), as suggested by network perspectives on personality (Costantini et al., 2015; Cramer et al., 2012; Schmittmann et al., 2013). From the network perspective, the observed correlations between personality items constitute direct reinforcement and inhibition processes between the behaviors. If someone likes people, they will be more inclined to meet strangers and will thus also tend to attend more social events. Positive experiences at these social events will then further increase this person's positive perception of other people, thus further enforcing their tendency to seek out new contacts. The cross-sectional trait level and longitudinal change correlations thus result from the same underlying network of reinforcement and inhibition across personality-related behaviors.

Another *mechanism* that could have contributed to these findings is response styles, in particular self-evaluation biases (Leising et al., 2015, 2020). The moderate positive cross-sectional correlations between Extraversion, Agreeableness, Openness, and Conscientiousness (but not Neuroticism) that were observed in this and previous correlated change studies (see Park et al., 2020 for a meta-analytic overview of cross-sectional Big Five correlations) can be attributed in part to the shared positive valence of the items and interindividual differences in socially desirable responding or self-evaluation (Leising et al., 2020). It is important to note that previous research on correlated change with broader and more balanced inventories such as the 60-item NEO Five-Factor Inventory (Costa & McCrae, 2004) still reported evidence of correlated change (see Allemand & Martin, 2016). This indicates that the current results are not merely an artifact of measurement error.

Contrary to the findings with respect to differential change, we did not find an effect of varying time intervals on correlated change patterns (see Table 2). Increasing time intervals generally have a negative effect on personality stability, as indicated by greater mean-level differences and lower rank-order stability with longer time intervals (Roberts et al., 2006; Roberts & DelVecchio, 2000). Within the MIDUS and HRS studies, change factor correlations were statistically equivalent across 9 and 18, and 4, 8, and 12 years, respectively. Independent of how much time has passed, correlated change patterns remained stable. Consequently, correlated change also remained similar to the initial baseline trait correlations. This characteristic explains why this finding was reported in the majority of previous studies on correlated change (e.g. Allemand et al., 2007, 2008; Klimstra et al., 2013; Mõttus et al., 2012), despite using varying time intervals. Of course, all studies on correlated change examined correlated change over several years, which might correspond to the time intervals required for the shared underlying processes to be observable as correlated change. Future studies could address whether these correlated change patterns can also be found in shorter time intervals, or how much time needs to pass until correlated change resembles the initial baseline trait-level correlations.

While differential change was unaffected by age, participants' age had a significant moderation effect on correlated personality change. Generally, correlated change increased with age, in particular for Conscientiousness from 25 to 50 years of age. Cross-sectional and change correlations in the older HRS sample were also particularly high. These findings may indicate a dedifferentiation in the personality change processes over time (cf. Allemand et al., 2008; Zimprich & Martin, 2010). This suggests that personality change in the various trait domains is more strongly interconnected in late adulthood. A common change factor was able to explain about 76% of the change in Openness, Conscientiousness, Extraversion, and Agreeableness in the older HRS sample, suggesting that the plasticity of personality traits might become more homogenous with age. Age-differentiation and dedifferentiation are aspects of structural change often examined in the field of cognitive development (e.g. Breit et al., 2020), where some researchers suggest that cognitive processes become increasingly independent in early life (i.e. differentiation; e.g. Li et al., 2004; but see also Salthouse, 2010; Tucker-Drob & Salthouse, 2008), and again more interrelated in old age as a consequence of cognitive decline (i.e. dedifferentiation; e.g. Hülür et al., 2015). Age-differentiation or dedifferentiation is generally examined through age-associated differences in the factor loadings on a common cognitive ability factor, or correlations between different cognitive domains (similar to our approach; e.g. Hartung et al., 2018).

In the field of personality development, differentiation and dedifferentiation processes are rarely studied, despite being highly informative for understanding interconnected processes of development. A recent study examining the structural stability of personality across the lifespan with a network approach also found substantial dedifferentiation in the personality trait structure with increasing age (Beck et al., 2019). As a consequence, the Big Five structure could not be retrieved in networks for older age groups. Issues of retrieving comparable or measurement invariant personality models across such broad age spans have also been reported in studies employing MGCFA (e.g. Olaru et al., 2018, 2019). Cross-cultural personality research suggests that the intercorrelations between personality items or traits may be dependent on the complexity of the society and the possibility to express a higher number of behavioral profiles (Smaldino et al., 2019). Physical limitations in old age, a decreasingly active lifestyle and a decreasing motivation for learning or improvement (e.g. Carstensen et al., 2003), might potentially also lead to such dedifferentiation or simplification processes in the context of personality structures. In this study, we focused on age-differentiation, which is often studied alongside ability-differentiation in research on cognitive abilities (e.g. Breit et al., 2020). The concept of ability-differentiation refers to differences in the structure or correlations of cognitive abilities depending on the ability mean-levels. Applying this concept to personality traits, future studies could examine structural differences in the personality traits (e.g. factor loading differences) or correlated change as a function of trait levels, and the interaction thereof with age.

### *Limitations and future directions*

Although the current study had multiple strengths, such as large longitudinal samples, tests of measurement invariance across the moderators, and the application of LSEM to maintain the continuous nature of age and to disentangle cross-sectional and longitudinal differences in correlated change, the present work has several limitations that may guide future research. Even though we combined two panel studies with broad age ranges, our findings are limited to adulthood. Because of the sample size requirements of latent variable modeling, we were only able to examine the effect of age on differential and correlated change from 25 to 80 years of age. To obtain a more comprehensive picture of differential and correlated change across age, a broader age range is desirable for future studies. We wonder whether we would find a personality change differentiation in childhood, similar to some studies on cognitive abilities (e.g. Li et al., 2004).

Both studies used the MIDI, which measures the Big Five with 25 single adjective items on a 4-point Likert scale. Whereas the MIDI is a comparatively long Big Five inventory in the context of panel studies, the scales are not balanced regarding item numbers. Even though the adjective markers were selected based on high factor main loadings and item-total correlations from longer adjective lists (Lachman & Weaver, 1997), the MIDI Big Five structures suffer from several residual correlations and cross-loadings (Zimprich et al., 2012). In addition, the 4-point Likert scale may be problematic in the context of personality development research, as the potential for change is restricted by the low number of response options. The reliance on only one personality inventory reduces the generalizability of our findings. In future studies, the comparison of comparable samples with varying personality inventories would be desirable in order to estimate the effect of items used on differential and correlated change, ideally including a facet level of personality as well.

We examined differential and correlated personality change with age and over time at a global level—only using age as a between-person indicator of development or life stages. The influence of specific life events or other contextual factors on differential and correlated change should be examined in future studies. For instance, it would be interesting to examine whether retirement or divorce affect several traits in tandem and thus correlated change. To do so, the models used in this study could be modified by combining the age moderation of LSEM with multigroup moderation approaches for categorical moderators (e.g. retirement). Similarly, the influence of other relevant continuous moderators could be included by extending the approaches used in this study. Future research should examine whether the increase in correlated change in old age is a result of cognitive decline (see, e.g. Klimstra et al., 2013) or of other age-related variables. To do so, the LSEM approach could be used simultaneously with two continuous moderator variables (e.g. age and cognitive abilities) to distinguish the effects of several moderators (see Hartung et al., 2018).

In this study, we focused on differential and correlated change in the context of personality traits. However, these two perspectives are relevant for the entire field of aging and lifespan development, and a more comprehensive picture including other personality-related variables is needed. For instance, examining correlated change between personality traits and health, cognitive functioning, or well-being across age can help understand how personality can contribute to healthy aging (e.g. Hill & Allemand, 2020). For instance, correlated change between Conscientiousness and health or other relevant outcomes can be studies across age using the procedures presented in this study. This approach can help identify potential bi-directionalities between

Conscientiousness and health (Mroczek et al., 2020) and at which age these effects are particularly strong.

## Conclusion

In this paper, we examined differential and correlated personality change as a function of varying time intervals ranging from 4 to 18 years, and participants' age at the first measurement occasion ranging from 25 to 80 years. We demonstrated how a combination of LSEM and longitudinal models can be used to examine complex developmental processes (see also Wagner et al., 2019). This approach allows researchers to differentiate between cross-sectional age differences (e.g. cohort effects) and longitudinal change within a common modeling framework without imposing restrictions on the age effects (e.g. linear, quadratic, and categorical). We found high levels of differential stability in adulthood and decreases in test–retest correlations with increasing time between measurements. Furthermore, we found that change processes in Extraversion, Openness, Agreeableness, and Conscientiousness were strongly related. Correlated change patterns corresponded roughly to the initial cross-sectional Big Five correlations but were on average slightly higher. Correlated change was stable across varying time intervals but increased with age. This trend indicates a dedifferentiation in the change processes of personality. This effect was strongest for change in Conscientiousness, which was only weakly related to the other change processes in young adulthood. From our view, the results presented here clearly illustrate the need to include correlated change as a significant informative perspective of lifespan development.

### Data accessibility statement

The analyses scripts and supplementary materials are available in an OSF repository (https://osf.io/x9u2j/). The data files used in this study can be obtained upon request and free of charge at https://www.icpsr.umich.edu/web/ICPSR/series/203 and https://hrs.isr.umich.edu/data-products

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### ORCID iD

Gabriel Olaru https://orcid.org/0000-0002-7430-7350

### Note

1. *Friendly* was dropped from the analysis because of strong cross-loadings on Agreeableness (see also Zimprich et al., 2012).

### Supplemental material

Supplemental material for this article is available online.

### References

Allemand, M., & Martin, M. (2016). On correlated change in personality. *European Psychologist*, *21*(4), 237–253. https://doi.org/10.1027/1016-9040/a000256

Allemand, M., Zimprich, D., & Hertzog, C. (2007). Cross-sectional age differences and longitudinal age changes of personality in middle adulthood and old age. *Journal of Personality*, *75*(2), 323–358. https://doi.org/10.1111/j.1467-6494.2006.00441.x

Allemand, M., Zimprich, D., & Martin, M. (2008). Long-term correlated change in personality traits in old age. *Psychology and Aging*, *23*(3), 545–557. https://doi.org/10.1037/a0013239

Anusic, I., & Schimmack, U. (2016). Stability and change of personality traits, self-esteem, and well-being: Introducing the meta-analytic stability and change model of retest correlations. *Journal of Personality and Social Psychology*, *110*(5), 766–781. https://doi.org/10.1037/pspp0000066

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, *81*(3), 261–272. https://doi.org/10.1037/0021-9010.81.3.261

Barry, T. R. (2014). The Midlife in the United States (MIDUS) series: A national longitudinal study of health and well-being. *Open Health Data*, *2*(1), pe3. https://doi.org/10.5334/ohd.ai

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling*, *13*(2), 186–203. https://doi.org/10.1207/s15328007sem1302_2

Beck, E. D., Condon, D. M., & Jackson, J. J. (2019, July 9). *Interindividual age differences in personality structure*. https://doi.org/10.31234/osf.io/857ev

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Billiet, J. B., & Davidov, E. (2008). Testing the stability of an acquiescence style factor behind two interrelated substantive variables in a panel design. *Sociological Methods and Research*, *36*(4), 542–562. https://doi.org/10.1177/0049124107313901

Borghuis, J., Denissen, J., Oberski, D., Sijtsma, K., Meeus, W., Branje, S., Koot, H. M., & Bleidorn, W. (2017). Big Five personality stability, change, and codevelopment across adolescence and early adulthood. *Journal of Personality and Social Psychology*, *113*(4), 641–657. https://doi.org/10.1037/pspp0000138

Borkenau, P., & Ostendorf, F. (1992). Social desirability scales as moderator and suppressor variables. *European Journal of Personality*, *6*(3), 199–214. https://doi.org/10.1002/per.2410060303

Breit, M., Brunner, M., & Preckel, F. (2020). General intelligence and specific cognitive abilities in adolescence: Tests of age differentiation, ability differentiation, and their interaction in two large samples. *Developmental Psychology*, *56*(2), 364. https://doi.org/10.1037/dev0000876

Briley, D. A., Harden, K. P., Bates, T. C., & Tucker-Drob, E. M. (2015). Nonparametric estimates of gene × environment interaction using local structural equation modeling. *Behavior Genetics*, *45*(5), 581–596. https://doi.org/10.1007/s10519-015-9732-8

Carstensen, L. L., Fung, H. H., & Charles, S. T. (2003). Socioemotional selectivity theory and the regulation of emotion in the second half of life. *Motivation and Emotion*, *27*(2), 103–123. https://doi.org/10.1023/A:1024569803230

Caspi, A., & Roberts, B. W. (2001). Personality development across the life course: The argument for change and continuity. *Psychological Inquiry*, *12*, 49–66. https://doi.org/10.1207/S15327965PLI1202_01

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, *56*, 453–484. https://doi.org/10.1146/annurev.psych.55.090902.141913

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *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

Chopik, W. J., & Kitayama, S. (2018). Personality change across the life span: Insights from a cross-cultural, longitudinal study. *Journal of Personality*, *86*(3), 508–521. https://doi.org/10.1111/jopy.12332

Cloninger, C. R., Przybeck, T. R., & Svrakic, D. M. (1991). The tridimensional personality questionnaire: US normative data. *Psychological Reports*, *69*(3), 1047–1057. https://doi.org/10.2466/pr0.1991.69.3.1047

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

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, *54*, 13–29. https://doi.org/10.1016/j.jrp.2014.07.003

Craik, F. I., & Bialystok, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends in Cognitive Sciences*, *10*(3), 131–138. https://doi.org/10.1016/j.tics.2006.01.007

Cramer, A. O., Van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., . . . & Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality*, *26*(4), 414–431. https://doi.org/10.1002/per.1866

Danner, D., Aichholzer, J., & Rammstedt, B. (2015). Acquiescence in personality questionnaires: Relevance, domain specificity, and stability. *Journal of Research in Personality*, *57*, 119–130. https://doi.org/10.1016/j.jrp.2015.05.004

DeYoung, C. G., & Gray, J. R. (2009). Personality neuroscience: Explaining individual differences in affect, behavior, and cognition. In P. J. Corr & G. Matthews (Eds.), *The Cambridge handbook of personality psychology* (pp. 323–346). Cambridge University Press. https://doi.org/10.1017/CBO9780511596544.023

Ferguson, C. J. (2010). A meta-analysis of normal and disordered personality across the life span. *Journal of Personality and Social Psychology*, *98*(4), 659–667. https://doi.org/10.1037/a0018770

Ferrer, E., & McArdle, J. J. (2010). Longitudinal modeling of developmental changes in psychological research. *Current Directions in Psychological Science*, *19*, 149–154. https://doi.org/10.1177/0963721410370300

Fraley, R. C., & Roberts, B. W. (2005). Patterns of continuity: A dynamic model for conceptualizing the stability of individual differences in psychological constructs across the life course. *Psychological Review*, *112*, 60–74. https://doi.org/10.1037/0033-295X.112.1.60

Haberecht, K., Schnuerer, I., Gaertner, B., John, U., & Freyer-Adam, J. (2015). The stability of social desirability: a latent change analysis. *Journal of personality, 83*(4), 404–412. https://doi.org/10.1111/jopy.12112

Hartung, J., Doebler, P., Schroeders, U., & Wilhelm, O. (2018). Dedifferentiation and differentiation of intelligence in adults across age and years of education. *Intelligence*, *69*, 37–49. https://doi.org/10.1016/j.intell.2018.04.003

Hennecke, M., Bleidorn, W., Denissen, J. J., & Wood, D. (2014). A three-part framework for self-regulated personality development across adulthood. *European Journal of Personality*, *28*(3), 289–299. https://doi.org/10.1002/per.1945

Hertzog, C., & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: An overview of methodological issues. *Psychology and Aging*, *18*, 639–657. https://doi.org/10.1037/0882-7974.18.4.639

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.

Hill, P. L., & Allemand, M. (Eds.). (2020). *Personality and healthy aging in adulthood: New directions and techniques*. Springer Nature.

Hofer, S. M., Flaherty, B. P., & Hoffman, L. (2006). Cross-sectional analysis of time dependent data: Mean-induced association in age-heterogeneous samples and an alternative method based on sequential narrow age-cohort samples. *Multivariate Behavioral Research*, *41*, 165–187. https://doi.org/10.1207/s15327906mbr4102_4

Holden, R. R., & Passey, J. (2010). Socially desirable responding in personality assessment: Not necessarily faking and not necessarily substance. *Personality and Individual Differences*, *49*(5), 446–450. https://doi.org/10.1016/j.paid.2010.04.015

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, *18*(3), 117–144. https://doi.org/10.1080/03610739208253916

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling* *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hudson, N. W., Briley, D. A., Chopik, W. J., & Derringer, J. (2019). You have to follow through: Attaining behavioral change goals predicts volitional personality change. *Journal of Personality and Social Psychology*, *117*(4), 839–857. https://doi.org/10.1037/pspp0000221

Hudson, N. W., & Fraley, R. C. (2015). Volitional personality trait change: Can people choose to change their personality traits? *Journal of Personality and Social Psychology*, *109*(3), 490–507. https://doi.org/10.1037/pspp0000021

Hülür, G., Ram, N., Willis, S. L., Schaie, K. W., & Gerstorf, D. (2015). Cognitive dedifferentiation with increasing age and proximity of death: Within-person evidence from the Seattle Longitudinal Study. *Psychology and Aging*, *30*(2), 311–323. https://doi.org/10.1037/a0039260

Hülür, G., Wilhelm, O., & Robitzsch, A. (2011). Intelligence differentiation in early childhood. *Journal of Individual Differences*, *32*, 170–179. https://doi.org/10.1027/1614-0001/a000049.

Jorgensen, T. D., Kite, B. A., Chen, P. Y., & Short, S. D. (2018). Permutation randomization methods for testing measurement equivalence and detecting differential item functioning in multiple-group confirmatory factor analysis. *Psychological Methods*, *23*(4), 708–728. https://doi.org/10.1037/met0000152

Juster, F. T., & Suzman, R. (1995). An overview of the Health and Retirement Study. *Journal of Human Resources*, *30*, 7–56. https://doi.org/10.2307/146277

Kandler, C., Kornadt, A. E., Hagemeyer, B., & Neyer, F. J. (2015). Patterns and sources of personality development in old age. *Journal of Personality and Social Psychology*, *109*(1), 175–191. https://doi.org/10.1037/pspp0000028

Klimstra, T. A., Bleidorn, W., Asendorpf, J. B., Van Aken, M. A., & Denissen, J. J. (2013). Correlated change of Big Five personality traits across the lifespan: A search for determinants. *Journal of Research in Personality*, *47*, 768–777. https://doi.org/10.1016/j.jrp.2013.08.004

Klimstra, T. A., Hale III, W. W., Raaijmakers, Q. A., Branje, S. J., & Meeus, W. H. (2009). Maturation of personality in adolescence. *Journal of Personality and Social Psychology*, *96*(4), 898–912. https://doi.org/10.1037/a0014746

Lachman, M. E., & Weaver, S. L. (1997). *The Midlife Development Inventory (MIDI) personality scales: Scale construction and scoring* (pp. 1–9). Brandeis University.

Leising, D., Scherbaum, S., Locke, K. D., & Zimmermann, J. (2015). A model of "substance" and "evaluation" in person judgments. *Journal of Research in Personality*, *57*, 61–71. https://doi.org/10.1016/j.jrp.2015.04.002

Leising, D., Vogel, D., Waller, V., & Zimmermann, J. (2020). Correlations between person-descriptive items are predictable from the product of their mid-point-centered social desirability values. *European Journal of Personality*, 0890207020962331. https://journals.sagepub.com/doi/full/10.1177/0890207020962331

Li, S. C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science*, *15*(3), 155–163. https://doi.org/10.1111/j.0956-7976.2004.01503003.x

Little, T. D. (2013). *Longitudinal structural equation modeling*. 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*, 847–861. https://doi.org/10.1037/a0024298.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*(1), 19–40. https://doi.org/10.1037/1082-989X.7.1.19

Martin, M., & Zimprich, D. (2003). Are changes in cognitive functioning in older adults related to changes in subjective complaints? *Experimental Aging Research*, *29*, 335–352. https://doi.org/10.1080/03610730303722

Mascherek, A., & Zimprich, D. (2011). Correlated change in memory complaints and memory performance across 12 years. *Psychology and Aging*, *26*, 884–889. https://doi.org/10.1037/a0023156

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, *60*, 577–605. https://doi.org/10.1146/annurev.psych.60.110707.163612

McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 139–175). American Psychological Association. https://doi.org/10.1037/10409-005

McCrae, R. R., Costa, P. T., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M. D., ... Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology*, *78*(1), 173–186. https://doi.org/10.1037/0022-3514.78.1.173

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

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. https://doi.org/10.1007/BF02294825

Meredith, W., & Horn, J. (2001). The role of factorial invariance in modeling growth and change. In L. M. Collins & A. G. Sayer (Eds.), *Decade of behavior. New methods for the analysis of change* (pp. 203–240). American Psychological Association. https://doi.org/10.1037/10409-007

Mõttus, R., Johnson, W., Starr, J. M., & Deary, I. J. (2012). Correlates of personality trait levels and their changes in very old age: The Lothian Birth Cohort 1921. *Journal of Research in Personality*, *46*(3), 271–278. https://doi.org/10.1016/j.jrp.2012.02.004

Mroczek, D. K., Weston, S. J., & Willroth, E. C. (2020). A lifespan perspective on the interconnections between personality, health, and optimal aging. In P. L. Hill, & M. Allemand, (Eds.), *Personality and healthy aging in adulthood* (pp. 191–202). Springer.

Mund, M., & Neyer, F. J. (2016). The winding paths of the lonesome cowboy: Evidence for mutual influences between personality, subjective health, and loneliness. *Journal of Personality*, *84*(5), 646–657. https://doi.org/10.1111/jopy.12188

Nye, C. D., Allemand, M., Gosling, S. D., Potter, J. & Roberts, B. W. (2016). Personality trait differences between young and middle-aged adults: Measurement artifacts or actual trends? *Journal of Personality*, *84*(4), 473–492. https://doi.org/10.1111/jopy.12173

Olaru, G., Schroeders, U., Hartung, J., & Wilhelm, O. (2019). Ant colony optimization and local weighted structural equation modeling. A tutorial on novel item and person sampling procedures for personality research. *European Journal of Personality*, *33*(3), 400–419. https://doi.org/10.1002/per.2195

Olaru, G., Schroeders, U., Wilhelm, O., & Ostendorf, F. (2018). A confirmatory examination of age-associated personality differences: Deriving age-related measurement-invariant solutions using ant colony optimization. *Journal of Personality*, *86*, 1037–1049. https://doi.org/10.1111/jopy.12373

Park, H. H., Wiernik, B. M., Oh, I.-S., Gonzalez-Mulé, E., Ones, D. S., & Lee, Y. (2020). Meta-analytic five-factor model personality intercorrelations: Eeny, meeny, miney, moe, how, which, why, and where to go. *Journal of Applied Psychology*, *105*(12), 1490–1529. https://doi.org/10.1037/apl0000476

Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 224–239). The Guilford Press.

Piedmont, R. L., McCrae, R. R., Riemann, R., & Angleitner, A. (2000). On the invalidity of validity scales: Evidence from self-reports and observer ratings in volunteer samples. *Journal of Personality and Social Psychology*, *78*(3), 582–593. https://doi.org/10.1037/0022-3514.78.3.582

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*(3), 354–373. https://doi.org/10.1037/a0029315

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*, 3–25. https://doi.org/10.1037//0033-2909.126.1.3

Roberts, B. W., & Helson, R. (1997). Changes in culture, changes in personality: The influence of individualism in a longitudinal study of women. *Journal of Personality and Social Psychology*, *72*(3), 641–651. https://doi.org/10.1037/0022-3514.72.3.641

Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin*, *143*(2), 117. https://doi.org/10.1037/bul0000088

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*, 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neosocioanalytic model of personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 11–39). Erlbaum.

Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In O. P. John, R.W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 375–398). Guilford.

Robitzsch, A. (2019). *sirt: Supplementary item response theory models. R package version 3.6-21*. https://CRAN.R-project.org/package = sirt

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*(2), 1–36.

Salthouse, T.A. (2010). *Major issues in cognitive aging*. Oxford University Press.

Schmittmann, V. D., Cramer, A. O., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, *31*(1), 43–53. https://doi.org/10.1016/j.newideapsych.2011.02.007

Schroeders, U., Schipolowski, S., & Wilhelm, O. (2015). Age-related changes in the mean and covariance structure of fluid and crystallized intelligence in childhood and adolescence. *Intelligence*, *48*, 15–29. https://doi.org/10.1016/j.intell.2014.10.006

Schwaba, T., & Bleidorn, W. (2019). Personality trait development across the transition to retirement. *Journal of Personality and Social Psychology*, *116*(4), 651–665. https://doi.org/10.1037/pspp0000179

Sliwinski, M. J., Hofer, S. M., & Hall, C. (2003). Correlated and coupled cognitive change in older adults with and without preclinical dementia. *Psychology and Aging*, *18*(4), 672–683. https://doi.org/10.1037/0882-7974.18.4.672

Smaldino, P. E., Lukaszewski, A., von Rueden, C., & Gurven, M. (2019). Niche diversity can explain cross-cultural differences in personality structure. *Nature Human Behaviour*, *3*(12), 1276–1283. https://doi.org/10.1038/s41562-019-0730-3

Soto, C. J., & John, O. P. (2012). Development of Big Five domains and facets in adulthood: Mean-level age trends and broadly versus narrowly acting mechanisms. *Journal of Personality*, *80*, 881–914. https://doi.org/10.1111/j.1467-6494.2011.00752.x

Specht, J., Bleidorn, W., Denissen, J. J. A., Hennecke, M., Hutteman, R., Kandler, C., Luhmann, M., Orth, U., Reitz, A. K., & Zimmermann, J. (2014). What drives adult personality development? A comparison of theoretical perspectives and empirical evidence. *European Journal of Personality, 28*(3), 216–230. https://doi.org/10.1002/per.1966

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

Stieger, M., Wepfer, S., Rüegger, D., Kowatsch, T., Roberts, B. W., & Allemand, M. (2020). Becoming more conscientious or more open to experience?

Effects of a two-week smartphone-based intervention for personality change. *European Journal of Personality*, *34*(3), 345–366. https://doi.org/10.1002/per.2267

Terracciano, A., Costa, P. T., Jr., & McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin*, *32*(8), 999–1009. https://doi.org/10.1177/0146167206288599

Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2003). Stability of self-esteem across the life span. *Journal of Personality and Social Psychology*, *84*(1), 205–220. https://doi.org/10.1037/0022-3514.84.1.205

Tucker-Drob, E. M., & Salthouse, T. A. (2008). Adult age trends in the relations among cognitive abilities. *Psychology and Aging*, *23*(2), 453–460. https://doi.org/10.1037/0882-7974.23.2.453

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

Weijters, B., Geuens, M., & Schillewaert, N. (2010). The individual consistency of acquiescence and extreme response style in self-report questionnaires. *Applied Psychological Measurement*, *34*(2), 105–121. https://doi.org/10.1177/0146621609338593

Wood, D., & Denissen, J. J. A. (2015). A functional perspective on personality trait development. In N. R. Branscombe & K. Reynolds (Eds.), *Psychology of change: Life contexts, experiences, and identities* (pp. 97–115). Psychology Press.

Wortman, J., Lucas, R. E., & Donnellan, M. B. (2012). Stability and change in the Big Five personality domains: Evidence from a longitudinal study of Australians. *Psychology and Aging*, *27*(4), 867–874. https://doi.org/10.1037/a0029322

Wrzus, C., & Roberts, B. W. (2017). Processes of personality development in adulthood: The TESSERA framework. *Personality and Social Psychology Review*, *21*(3), 253–277. https://doi.org/10.1177/1088868316652279

Zimprich, D., Allemand, M., & Lachman, M. E. (2012). Factorial structure and age-related psychometrics of the MIDUS personality adjective items across the life span. *Psychological Assessment*, *24*, 173–186. https://doi.org/10.1037/a0025265

Zimprich, D., & Martin, M. (2002). Can longitudinal changes in processing speed explain longitudinal age changes in fluid intelligence? *Psychology and Aging*, *17*, 690–695. https://doi.org/10.1037/0882-7974.17.4.690

Zimprich, D., & Martin, M. (2010). Differentiation-dedifferentiation as a guiding principle for the analysis of lifespan development. In A. Kruse (Ed.), *Leben im Alter. Eigen- und Mitverantwortlichkeit in Gesellschaft, Kultur und Politik* (pp. 29–37). Akademische Verlagsgesellschaft Aka.