Unravelling the interplay between genetic and environmental contributions in the unfolding of personality differences from early adolescence to young adulthood

Unraveling the Interplay between Genetic and Environmental Contributions in the Unfolding of Personality Differences from Early Adolescence to Young Adulthood

Christian Kandler[1], Trine Waaktaar[2], René Mõttus[3], Rainer Riemann[4], & Svenn Torgersen[2]

Uncorrected preprint version of a paper accepted for publication. Please cite it as follows: Kandler, C., Waaktaar, T., Mõttus, R., Riemann, R., & Torgersen, S. (in press). Unraveling the interplay between genetic and environmental contributions in the unfolding of personality differences from early adolescence to young adulthood. *European Journal of Personality*, http://dx.doi.org/10.1002/per.2189

[1] University of Bremen, Bremen, & MSB Medical School Berlin, Berlin, Germany

[2] University of Oslo, Oslo, Norway

[3] University of Edinburgh, Edinburgh, UK, and University of Tartu, Tartu, Estonia

[4] Bielefeld University, Bielefeld, Germany

Correspondence concerning this proposal should be addressed to Christian Kandler, Department of Psychology, University of Bremen, Personality Psychology and Psychological Assessment, Grazer Strasse 2c, 28359 Bremen, Germany

Email: ckandler@uni-bremen.de

Phone: +49 421 218-68770

Funding information: The TwinLife study is supported by a grant from the German Research Foundation awarded to Martin Diewald (DI 759/11-1), Rainer Riemann (RI 595/8-1), and Frank M. Spinath (SP 610/6-1); the NTSY was funded by grants 170089, 213722 and 213760 from the Research Council of Norway. During the time of writing this paper, Christian Kandler's research was supported by grants from the German Research Foundation (KA-4088/2-1 and KA-4088/2-2).

Abstract

In two studies, we examined the genetic and environmental sources of the unfolding of personality trait differences from childhood to emerging adulthood. Using self-reports from over 3,000 representative German twin pairs of three birth cohorts, we could replicate previous findings on the primary role of genetic sources accounting for the unfolding of inter-individual differences in personality traits and stabilizing trait differences during adolescence. More specifically, the genetic variance increased between early (ages 10-12) and late adolescence (age 16-18) and stabilized between late adolescence and young adulthood (ages 21-25). This trend could be confirmed in a second three-wave longitudinal study of adolescents' personality self-reports and parent ratings from about 1,400 Norwegian twin families (average ages between 15 and 20). Moreover, the longitudinal study extended previous research and provided evidence for increasing genetic differences being primarily due to accumulation of novel genetic influences instead of an amplification of initial genetic variation. This is in line with cumulative interaction effects between twins' correlated genetic makeups and environmental circumstances shared by adolescent twins reared together. In other words, nature × nurture interactions rather than transactions can account for increases in genetic variance and thus personality variance during adolescence.

Keywords

Genotype × environment transaction and interaction; Personality differences; Twin study; Genetic and environmental variance; Adolescence

Introduction

Whereas personality development research has paid considerable attention to age trends in mean levels of traits and their rank-order stability (Roberts & DelVecchio, 2000; Roberts, Walton, & Viechtbauer, 2006; Specht et al., 2014), comparatively few studies have investigated how and why the magnitude of inter-individual differences in traits changes across the lifespan (Mõttus, Allik, Hřebíčková, Kööts-Ausmees, & Realo, 2016; Mõttus, Soto, & Slobodskaya, 2017). We examine which developmental mechanisms can explain the unfolding of trait differences from childhood to emerging adulthood and which account for the stabilization of trait variance during late adolescence. We first review the literature on age trends in the magnitude of inter-individual personality differences from childhood to adulthood. We then compare different theoretical explanations for reported findings, identify gaps in them and argue that twin studies are suitable for refining and testing the explanations. Using two samples from genetically informative cohort-sequential and longitudinal twin projects from Germany and Norway, we investigate genetic and environmental contributions to the unfolding of inter-individual differences in Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness between early adolescence and emerging adulthood (from 10 to 25 years of age).

Age Trends in Inter-Individual Personality Trait Variance from Childhood to Adulthood

Using parent reports of two large age- and gender-balanced samples of youth between 2 and 20 years, Mõttus and colleagues (2017) found strong evidence for increasing individual differences in all Big Five personality trait scores (except Extraversion) from early childhood to early adolescence. This pattern was replicated in parent ratings and self-reports of American children primarily between 8-18 years of age (Mõttus, Briley et al., in press). Both studies indicated that the magnitude of trait differences tend to plateau in mid-adolescence. Consistently, Mõttus et al. (2016) found no systematic differences in any personality trait and facet variance between late adolescents (ages 16-20) and young adults (ages 21-25) from Russia and Estonia. Moreover, there appear to be no systematic trends of increasing or decreasing magnitude of personality trait differences in adulthood (Allemand, Zimprich, & Hendriks, 2008; Loehlin & Martin, 2001; Soto, John, Gosling, & Potter, 2011). For example, Mõttus et al. (2016) found no consistent differences between young adults and middle-aged adults in the variance of self- and informant-rated traits in Estonia, the Czech Republic, and Russia. If anything, there was a decrease in the variance of Conscientiousness and related

facets (e.g., Dutifulness), but no systematic increase in the variance of any trait with age. A longitudinal multi-rater study of German twins, modelling personality variables as latent trait scores based on self- and informant reports (Kandler et al., 2010), suggested a similar decimation of inter-individual variance in Conscientiousness in young adults.

Developmental Explanations for Age Differences in Inter-Individual Trait Differences

As one possible explanation for such age trends in personality trait variance, inter-individual differences in unsystematic experiences or opportunities that occur by chance can act to increase trait variance. For example, although genetically identical twins (i.e., same genes) share the same womb (i.e., same environment) before birth, their prenatal positions are due to chance and can result in differences in blood supply and metabolism generating unique growth processes and outcomes (e.g., unique fingerprints). Results of such highly idiosyncratic growth processes can amplify over time and contribute to the increasing magnitude of inter-individual differences (Molenaar, Boomsma, & Dolan, 1993; Molenaar, Huizinga, & Nesselroade, 2003). Moreover, different random experiences and events may *accumulate* intra-individually with age and interact with each other in a very complex and highly idiosyncratic way (Plomin & Daniels, 2011; Turkheimer & Waldron, 2000). For example, depending on the individual situation and life circumstances, a workplace promotion may be experienced as a positive development by one person, associated with financial and status improvements, but as a more salient negative change for other persons, associated with separation from the family for a longer time. We subsume these environmental amplification and accumulation explanations of inter-individual trait differences as *environmental individualization hypotheses*.[5]

In line with the environmental individualization hypotheses, younger people tend to report more life events (e.g., moves, several graduations, and start an own family) than do middle-aged adults (Arnett, 2000; Kandler, Bleidorn, Riemann, Angleitner, & Spinath, 2012). Even though many of these events are unlikely to happen to people randomly (independently of their own behaviour and thereby personality), more interactions between differential social demands to deal with and adapt to may occur in younger ages. This may account for lower environmental stability of personality variance in youth compared to adult

[5] Note that unsystematic experiences can also decrease the magnitude of variance if most random experiences contribute to average rather than extreme trait levels. In fact, even the accumulation of unsystematic experiences that contribute to extreme trait levels can act to balance rather than increase trait differences (as in central limit theorem; see Mõttus, Allerhand, & Johnson, 2017).

ages. Cross-sequential twin studies (Kandler et al., 2010; Viken, Rose, Kaprio, & Koskenvuo, 1994) have revealed that the stabilization of trait variance in young adulthood may primarily be due to the stabilization of environmental sources (not shared by twins), whereas genetic variance tends to act as immutable basis.

Although most twin studies have not considered increasing or decreasing personality trait variance, they have shown that both environmental and genetic sources contribute to inter-individual differences in trait change in younger ages (Bratko & Butković, 2007; Gillespie, Evans, Wright, & Martin, 2004; Hopwood et al., 2011; Spengler, Gottschling, & Spinath, 2012). One recent study reported that the increase of personality trait variance from childhood through mid-adolescence was primarily attributable to increasing genetic differences (Mõttus, Briley et al., in press). That is, the emergence of personality differences appeared to be genetically driven rather than due to amplifying effects or accumulation of life experiences.

The first two decades of life are a time of expansion of individual capacities and innate basic tendencies, characterized by physiological, cognitive, and socio-emotional maturation (McAdams, 2015; McCrae & Costa, 2008; Piaget, 1970). Similar to body growth, systematic mean-level trends in specific personality traits may be due to an evolved, genetically driven *intrinsic maturation*, with different cultures and genetically related species showing similar normative age trends (McCrae et al., 1999; McCrae et al., 2000; Weiss & King, 2015). In the same vein, according to the *genetic maturation hypothesis,* the expansions of individuals' deviations from normative trends may be due to an unfolding of individuals' genetic endowments (i.e., genotypes; Kandler, 2012; Mõttus, 2017), just as individual differences in height take time to emerge in their adult-like magnitude and degree of individual differences (Tanner, Whitehouse, & Takaishi, 1966).

The degree to which genetically driven maturation underlies personality development, however, may depend on opportunities and the limits of environmental resources (Scarr, 1992, 1993), similarly to how unfolding of height differences depends on the availability of adequate nutrition (Johnson, 2010). This phenomenon is known as *genotype × environment interaction* (Eaves, Last, Martin, & Jinks, 1977; Plomin, DeFries, & Loehlin, 1977). For instance, total and genetic variance in negative emotionality have been found to be lower for 17-year-old adolescents who experienced lower levels of parental regard and higher levels of parental conflict (Krueger, South, Johnson, & Iacono, 2008). Environmental

circumstances may systematically change with age, providing different opportunities for genetic expression. If so, estimated increases in the genetic component of trait variance may result from *cumulative interaction effects between genotypes and environments* shared by individuals raised in the same family. Those interaction effects act as a function of the genetic relatedness of individuals and thus appear as estimates of the genetic variance, if not directly estimated in behavior genetic studies (Briley, Livengood, & Derringer, 2018).

Complicating things further, individuals can play an active role in their trait development (McAdams & Olson, 2010; McAdams, 2015). Since birth, babies' differential behavioral tendencies stimulate differential responses from their social environment that could act to reinforce or reduce the preexisting tendencies. As they grow, children gain more freedom and autonomy from parents and teachers, which may come along with increasing opportunities to actively shape and regulate their own development: People can be attracted to, create, or invest in niches and/or social roles that are consistent with their preexisting traits and would allow them to express themselves; they rather avoid contexts that are inconsistent with their predispositions and tend to change environments so that they fit better with their preexisting tendencies. This active and evocative role can explain both increasing stability of personality differences (what is known as the niche-picking principle of personality stabilization; Roberts & Nickel, 2017) and the accentuating of their variance (the corresponsive principle of personality development; Caspi, Roberts, & Shiner, 2005). Moreover, as the initial impetus may partly come from the individuals' genetic makeup, such increasing and stabilizing of personality differences can reflect underlying genetic differences unfolding *and* crystallizing over time (Kandler & Zapko-Willmes, 2017). Since the nonrandom exposure of individuals' genotypes to certain environments is known as *genotype–environment correlation* or *transaction* (Briley et al., 2018; Scarr & McCartney, 1983), we call this the *genotype × environment transaction hypothesis*.

Although a direct examination of genotype × environment transactions is difficult, phenotypic longitudinal studies have provided indirect evidence in terms of person × environment transactions (see Roberts & Nickel, 2017, for an overview). For example, Denissen, Ulferts, Lüdtke, Muck, and Gerstorf (2014) reported transactional effects between job environment and extraversion as well as openness in a 5-year longitudinal study of job beginners, job stayers, and job changers. Based on a 16-year longitudinal study of primarily adult participants, Jeronimus, Riese, Sanderman, and Ormel (2014) found that neuroticism

and negative life experiences showed bidirectional and persistent reinforcement. But most evidence for corresponsive person × environment transaction stems from investigations of young and middle-aged adults (e.g., Lüdtke, Roberts, Trautwein, & Nagy, 2011; Roberts, Caspi, & Moffitt, 2003; Roberts & Robins, 2004; Roberts, Walton, Bogg, & Caspi, 2006; Zimmermann & Neyer, 2013), whereas there is little evidence for increasing personality differences during adulthood (Allemand et al., 2008; Loehlin & Martin, 2001; Mõttus et al., 2016; Soto et al., 2011). These conflicting results might be resolved by developmental mechanisms that counterbalance person (or genotype) × environment transactions in adult age. Indeed, it is hard to see a vicious circle between ever-increasing neuroticism and ever more negative life events lasting in perpetuity.

Another well-established principle of personality development, the *maturity principle* (Caspi et al., 2005; Roberts & Nickel, 2017), may counterbalance the corresponsive principle. Originally conceptualized to account for systematic mean-level increases in Agreeableness, Emotional Stability, Conscientiousness, and Social Assertiveness, it postulates socialization processes that may not only push mean levels toward social maturity but also increase uniformity (Mõttus et al., 2016). Particularly during the transition to adulthood, when young people leave their parental home, start an apprenticeship or job, and invest into their own family, social demands may set strong normative standards for socially functional and successful behavior, and may decrease behavioral differences among individuals (Denissen, van Aken, Penke, & Wood, 2013; Mõttus et al., 2017); we call this the *environmental normalization hypothesis*.

These different explanations (see Table 1 for an overview) for the observed trends in the magnitude of inter-individual personality differences during development are not mutually exclusive. Each can contribute to the unfolding, decimation, or stabilization of the extent of personality differences with age. Based on observed personality scores alone, it is hard to tell the explanations apart, but studies on the personality similarity and differences between twin siblings of different ages can help to disentangle the relative contributions of genetic and environmental sources to the unfolding of personality differences.

Unraveling the Sources of Age Trends in Personality Trait Variance with Twin Studies

Studying twins reared together allows disentangling of the net genetic component of inter-individual differences (i.e., the heritability h^2) from variation due to environmental

sources.[6] If genetic differences contribute to the variance in a trait, genetically identical monozygotic (MZ) twins should be more similar in it than fraternal or dizygotic (DZ) twins who share about 50% of sharing segregating gene variants. A larger MZ twin similarity compared to the DZ twin similarity ($r_{MZ} > r_{DZ}$) has been found for nearly all human traits including personality characteristics (Polderman et al., 2015). On average across age groups, about 50% of variance in personality traits is due to genetic differences (i.e., $h^2 \approx .50$). Since twins raised in the same household may also share age-related and other common environmental influences, their resemblance may also result from shared environmental influences; this is the case when DZ twins are more similar in a trait than is expected from their genetic similarity. However, twin studies, but also other behavioral genetic designs, have mostly yielded negligible estimates of shared environmental influences (see Johnson, Vernon, & Feiler, 2008, and Vukasović & Bratko, 2015, for meta-analyses).

The estimation of the genetic variance in traits based on the design of twins reared together often relies on the assumption that gene variants additively (i.e., independently of each other) contribute to observed trait differences (i.e., additive genetic component a^2). However, genetic variance can also result from interactions between gene variants (i.e., nonadditive genetic component na^2) either within a gene locus (i.e., allelic dominance) or between gene loci (i.e., emergenesis), where the effect of one gene variant on a specific trait depends on the presence of one or more other modifying gene variants (i.e., $h^2 = a^2 + na^2$). Whereas genetically identical MZ twins share 100% of those nonadditive genetic sources, DZ twins have a 25% probability of sharing allelic dominance effects and they share polygenic gene × gene interaction effects with a probability close to zero (Lykken, 1982, 2006). Accordingly, nonadditive genetic contributions to the variance are indicated when MZ twin correlations are more than twice as large as DZ twin correlations ($r_{MZ} > 2 \times r_{DZ}$). Strong evidence for a significant nonadditive genetic source – primarily emergenesis – accounting for about a half of the genetic differences in personality traits has been confirmed by several genetically informative studies, including classic twin designs and beyond (e.g., Pilia et al., 2006; Plomin, Corley, Caspi, Fulker, & DeFries, 1998; Vukasović & Bratko, 2015).[7]

---

[6] Please note that estimates of genetic components or heritability of traits does not imply genetic determinism (see also Visscher, Hill, and Wray, 2008, for other misconceptions). Genetic variance and thus heritability or environmental variance components are population-based parameters that can vary across samples, time, and age.

[7] We do not list all limitations of classic twin designs, such as the no-assortative-mating assumption or the equal-environment assumption, because other studies including extended twin family designs

Because MZ twins' dissimilarity can only result from environmental influences, it quantifies the contributions of individualizing experiences that are independent of genetic influences and thereby make siblings reared together different from one another (i.e., nonshared environmental component e^2). The comparison of the MZ twin siblings' trait covariance (COV_{MZ}) with the total trait variance (VAR_{Trait}) allows us to estimate this component: $e^2 = (VAR_{Trait} - COV_{MZ})/VAR_{Trait}$. If only individualizing influences account for the increase in trait variance with age, twin covariance should not increase, resulting in declining twin correlations and heritability estimates with age (see Figure 1A and supplementary Table S1 for more details regarding the scenarios shown in Figure 1 and its consequences for twin correlations and estimates of genetic and environmental components). This pattern is in line with the findings of the meta-analysis by Briley & Tucker-Drob (2014), which suggested declining importance of genetic variance and an accumulation of individualizing experiences (*environmental individualization hypotheses*) either occurring at random or interacting (but not correlating or transacting) with an individual's unique genetic makeup (see Table 1). With increasing age, individuals become more independent from their parents and other caring family members, having more opportunities for individually unique experiences.

A more recent meta-analysis (Kandler & Papendick, 2017), however, revealed an *increasing* heritability from childhood to emerging adulthood and declines thereafter. This can be expected in cases of a) unfolding of individuals' genetic endowments via genetic maturation (Kandler, 2012a), b) increasingly active person-environment transactions (Scarr & McCartney, 1983), or c) cumulative interactions between the genotype and environmental circumstances shared by twins (Purcell, 2002). If increasing trait variance comes along with increasing genetic variance (*genetic unfolding hypotheses*, see Table 1), MZ twin covariance should increase to the same extent, whereas DZ twin covariance should increase proportionally to their genetic relatedness. This results in overall increases in twin correlations and heritability estimates with age (see Figure 1b and Table S1).

In the first systematic examination of the sources of increasing Big Five personality trait variance, Mõttus, Briley, and colleagues (in press) used self-reports and parent ratings

---

or different behavior genetic designs not limited to these assumptions support the primary results of twin studies regarding the net contributions of genetic and environmental sources to personality variance. In the current study, we primarily address the assumption of the independence of genetic and environmental influences and how estimates of genetic and environmental variance components from longitudinal and age-cohort twin studies can help to shed more light on the role of genotype × environment interplay in personality development.

from a population-based sample of over 2,500 twins from the Texas Twin Project, mostly in ages 8-18 years (Harden, Tucker-Drob, & Tackett, 2013). Across the Big Five traits, they found support for increases in genetic variance primarily attributable to increasing nonadditive genetic contributions (assuming allelic dominance effects for their modeling strategy). The primary increase of the nonadditive genetic component could be due to novel activations of gene variants that interact with other variants during developmental transitions (e.g., puberty), amplifying transactions between behavioral differences due to nonadditive genetic sources and environmental influences, or cumulative interaction effects between nonadditive genetic sources and environmental circumstances shared by twins. In this case, MZ twin correlations and heritability estimates would increase, whereas DZ twin correlations and their genetic correlation would decline in size (see Figure 1C and Table S1).[8]

Study 1: A Twin Study on the Sources of Age Trends in Personality Trait Variance from Late Childhood to Young Adulthood

In a first study, we aimed to replicate the findings by Mõttus, Briley, and colleagues (in press), analyzing personality self-reports from over 3,000 representative German twin pairs of three birth cohorts taken from the TwinLife study (Hahn et al., 2016). Specifically, we expected no significant differences in the magnitude of genetic variance in personality traits between the two older cohorts (late adolescence: ages 16-18; young adulthood: ages 21-25), but significantly smaller genetic contributions to trait variance in the youngest cohort (late childhood: ages 10-12).[9] Increasing genetic variance would be in line with three of the explanations introduced above: Genetic unfolding via genetic maturation, whereby genes gradually shine through all other influences on personality, an accumulation of genotype × environment interactions, whereby genotypes interact with environmental circumstances

[8] It must be noted, however, that what appears as nonadditive genetic variance may also result from processes occurring at the level of phenotype, rather than only due to allelic interactions. For example, when the phenotype consists of multiple causally interconnected components, even additive-only genetic influences on these components can yield nonadditive variance in the aggregate of these components (Mõttus & Allerhand, 2018). Likewise, when person × environment transactions are driven by only a subset of characteristics, they may result in genetically more similar individuals experiencing exponentially more similar environments, leading to nonadditive-like genetic variance (Mõttus, Briley et al., in press).

[9] Please note that study 1 is not a direct replication study of the study by Mõttus, Briley et al. (in press). Both studies slightly differ in age range (4-21 vs. 10-25) and TwinLife provided data from three clearly separable birth cohorts. Thus, in the current study, the age effects were examined by comparing age cohorts, rather than testing the moderating effects of age directly in the model.

shared by twins, and amplifying effects of genotype × environment transactions (see Table 1 for the *genetic unfolding* hypotheses 3a/b, 4a/b, and 5a/b).

According to Mõttus, Briley et al. (in press), who identified nonadditive genetic factors as primary sources of the increase in the genetic component from childhood to adolescence, we expected decreasing genetic correlations between DZ twin siblings across cohorts (*Genetic unfolding* hypotheses 3b, 4b, and 5b). We also tested for significant differences in the magnitude of the environmental component across age groups, with increasing environmental variance being in line with the *environmental individualization* hypotheses: Amplification of initial environmental differences or accumulation of novel life experiences independent of genetic influences, or accumulation of genotype × environment interactions, whereby genotypes interact with environmental circumstances *not* shared by twins (Table 1: hypotheses 1, 2, and 6). Decreasing environmental variance would support the *environmental normalization* hypothesis, whereby common social demands act to decrease trait differences (Table 1: hypothesis 7). We expected the patterns to be similar across traits.

Method

Participants

This study was based on twin data from the first wave of the TwinLife project (Hahn et al., 2016), which is an ongoing genetically informative, cohort-sequential extended twin family study of genetic and social causes of life chances and social inequality. The sample is representative for German families regarding income, education, and occupational status (the scientific use file is available at https://dbk.gesis.org/DBKSearch/SDesc2.asp?no=6701). Currently, it contains data from the first face-to-face survey for the full sample including about 4,000 families with same-sex twin pairs from four different birth cohorts (Cohort 1: 2009/2010; Cohort 2: 2003/2004; Cohort 3: 1997/1998; Cohort 4: 1990-1993). Nearly all twins (> 99%) from cohorts 2, 3, and 4 provided personality self-ratings (see Table 2 for an overview on descriptive sample statistics and Hahn et al., 2016, for more details on the TwinLife project including recruitment procedure, zygosity determination, and representativeness). Missing values (< 1%) have been replaced by a regression based on expectation maximization procedures (Little & Rubin, 2002).

Measures

Twins provided self-ratings on the 16-item Big Five Inventory (BFI-S; Hahn, Gottschling, & Spinath, 2012). This measure shows acceptable levels of psychometric quality. It captures

the Big Five personality domains reasonably well. Respondents were asked to rate themselves on the 16 items (4 items capturing Openness and 3 items for the other domains) with a 7-point Likert scale (1 "does not apply to me at all" to 7 "applies to me perfectly").

Similar to Mõttus, Briley and colleagues (in press), we prepared the data for main analyses. First, all items were corrected for acquiescence and extreme responding by centralizing each individual score on the individual-specific mean and standard deviation of responses to pairs of BFI-S items with opposite implications for personality traits (e.g., "I see myself as someone who gets nervous easily" and "…is relaxed, handles stress well"; see also Soto, John, Gosling, & Potter, 2008, for more details). In the current study, age and sex differences in acquiescence and extreme responding were marginal: ρ's < |.15|. However, in line with previous studies (Mõttus, Briley et al., in press; Soto et al., 2008), the variance of acquiescence was larger for the younger cohort (average age 11) compared to the two older cohorts (mean ages 17 and 23): VAR_{11} = .44 versus VAR_{17} = .26 and VAR_{23} = .28. Variance in extreme responding also tended to be larger for the younger cohort: VAR_{11} = .33 versus VAR_{17} = .22 and VAR_{23} = .25.

After correction for acquiescence and extreme responding, principal component analyses with varimax rotation yielded a clear five-factor structure that accounted for about 60% of the variance and all items loaded highest on their respective personality dimension, except for the youngest twin cohort in which factors were less structurally independent (see supplementary Table S2). After item recoding and z-standardization across all cohorts[10], we calculated Cronbach's α as index of the lower bound of internal consistency. The average internal consistency of measures across all Big Five traits was lower for the youngest twin cohort (average α_{11} = .50; α_{11} = .54 for Neuroticism, α_{11} = .56 for Extraversion, α_{11} = .54 for Openness, α_{11} = .37 for Agreeableness, and α_{11} = .55 for Conscientiousness) compared to those values for the 17 years old cohort and young adult twins. For the older cohorts the alphas were largely consistent: average α = .65 for both cohorts, α_{17} = .63 and α_{23} = .62 for Neuroticism, α_{17} = .80 and α_{23} = .78 for Extraversion, α_{17} = .62 and α_{23} = .65 for Openness, α_{17} = .56 and α_{23} = .53 for Agreeableness, as well as α_{17} = .68 and α_{23} = .65 for

[10] We did not correct for age differences within cohorts, because age ranges and effects were rather small, ranging between β = -.06 and β = .10 and statistical significance (p < .01) was not consistent across twin i and co-twin j subsamples. Across cohorts, however, we found significant but small linear age effects on two Conscientiousness items (β: .10 to .16; p < .001), all Openness items (β: -.19 to -.08; p < .001), and one Neuroticism item (β: .11 to .12; p < .001) that were statistically significant for both twin i and co-twin j subsamples.

Conscientiousness. This pointed to a lack of measurement invariance (MI) across birth cohorts that had to be taken into account in the main analyses.

Analyses

All initial and preparatory analyses were conducted with the statistical software package IBM SPSS 21.0. The main analyses were run with the add-on software IBM SPSS AMOS 21.0 (Arbuckle, 2012). The commented syntax for data preparation and initial analyses as well as AMOS scripts for structural equation model (SEM) analyses described in the following sections are retrievable from the Open Science Framework (https://osf.io/wk9rp/; see folder Study 1).

As internal consistency and factorial validity were lower for BFI personality scores of younger raters, twin correlations could be artificially lower for younger twins compared to those between older twins leading to artificially lower estimates of genetic components for younger twins due to lack of MI. This had to be taken into account by use of an appropriate latent variable measurement approach (see Figure 2). This SEM allows estimations of latent trait score variances controlled for error variance based on the assumption of τ-congeneric item scores (i.e., $\lambda = 1$ for those item indicators of a personality trait with the highest loading and free estimates for all other factor loadings, unequal and independent error variances). It also controls for sex differences, which typically arise during puberty (Soto et al., 2011).

The latent trait score variances, sex differences, and error variances were allowed to vary between the three age cohorts *x*. In this regard, the SEM enabled tests for metric and strict MI across cohorts. An assumption of metric MI would require that factor loadings are invariant across all cohorts (i.e., $\lambda_{1,x=11} = \lambda_{1,x=17} = \lambda_{1,x=23}$, $\lambda_{2,x=11} = \lambda_{2,x=17} = \lambda_{2,x=23}$, and $\lambda_{3,x=11} = \lambda_{3,x=17} = \lambda_{3,x=23}$; cf. Figure 2). Metric MI is the basic prerequisite to assume that same trait constructs have been measured in each cohort. Strict MI would additionally require that error variances and error correlations within same items between twin siblings are invariant among cohorts: Metric MI plus $VAR(\varepsilon_{1,x=11}) = VAR(\varepsilon_{1,x=17}) = VAR(\varepsilon_{1,x=23})$, $VAR(\varepsilon_{2,x=11}) = VAR(\varepsilon_{2,x=17}) = VAR(\varepsilon_{2,x=23})$, and $VAR(\varepsilon_{3,x=11}) = VAR(\varepsilon_{3,x=17}) = VAR(\varepsilon_{3,x=23})$, as well as $m_{1,x=11} = m_{1,x=17} = m_{1,x=23}$, $m_{2,x=11} = m_{2,x=17} = m_{2,x=23}$, and $m_{3,x=11} = m_{3,x=17} = m_{3,x=23}$ (cf. Figure 2). We did not test for the equality of means and intercepts across cohorts (scalar MI), because no mean-level comparisons were planned.

Variance-covariance matrices were fitted to the models using maximum likelihood procedures. The overall model fit was evaluated with two common criteria: The Root Mean

Square Error of Approximation (RMSEA) < .08 and the Comparative Fit Index (CFI) > .90 indicate an acceptable model fit (Hu & Bentler, 1999; Steiger, 1990). By comparing of the model fit of metric MI and strict MI models with the model fit of the unconstrained model, we tested for the equality of the constrained parameters under metric and strict MI conditions. For assessing relative model fit of nested models, we compared the 90% RMSEA confidence intervals (CIs) for all models (MacCallum, Browne, & Sugawara, 1996). Since the RMSEA is virtually independent of sample size, overlapping 90% RMSEA CIs provides an adequate test for MI. As a further descriptive criterion to identify lack of invariance, we used the ΔCFI < .01 criterion: A constrained model should not show a decrease in the CFI value larger than .01 (Cheung & Rensvold, 2002).

Based on the best fitting measurement models, we examined our hypotheses with the χ^2-difference test to detect differences between the cohorts with sufficient statistical power. Our expectations regarding a genetic unfolding during adolescence were tested via two model comparisons – whether models with $g_{x=11} = g_{x=17}$ (expecting a difference) and $g_{x=17} = g_{x=23}$ (expecting no difference) fitted significantly worse compared to a model allowing for varying genetic differences among cohorts. The model further allowed an examination of whether additive, nonadditive, or both genetic sources were involved in the increase of the genetic variance. In this respect, we tested for the invariance of DZ twins' genetic correlations ($\sigma_{x=11} = \sigma_{x=17} = \sigma_{x=23}$) across cohorts. Genetic correlations lower than .50 would indicate nonadditive genetic sources. Finally, we tested for the invariance of environmental contributions ($e_{x=11} = e_{x=17} = e_{x=23}$).

Results and Discussion

As shown in Table 3, initial (unconstrained) models for all Big Five traits provided at least acceptable model fit (RMSEAs < .03 and CFIs > .92). Taking all criteria for model comparisons into account (i.e., overlapping 90% RMSEA CIs and ΔCFI < .01), metric MI across cohorts could be established for all Big Five trait scores, whereas strict MI could be supported for none of them. The latter was primarily due to varying size in residual variances between the youngest and the two older cohorts (consistent with the differences in internal consistency). Based on models assuming only metric MI across cohorts[11], latent biometric model parameter estimates for all Big Five traits are shown in Table 4 (see also Table S3 for

[11] We also tested for MI across sexes: Metric MI could be supported for all Big Five traits (ΔCFI ≤ .003).

all model parameter estimates including sex effects, error variances, factor loadings, and item residual correlations with same items between twin siblings).

In line with our expectation (genetic unfolding during adolescence), the genetic effect on inter-individual trait differences was larger for late adolescence (age 17) compared to the youngest cohort (age 11). The heritability – estimated as $g^2/(g^2 + e^2)$ – increased on average from .37 (range: .30 – .46) to .48 (range: .36 – .58). Model tests indicated a significant difference for all Big Five personality traits, except for Neuroticism and Agreeableness (see model tests in Table 4). Also consistent with our expectation, there were no significant differences between the two older cohorts (age 17 vs. age 23), except in Neuroticism. The heritability estimates for age 23 (on average: .50; range: .39 – .58) were comparable to those of age 17.

For Neuroticism, we found substantial sex differences (females showed higher trait scores) that amplified between age 11 and 17 (from .226 to .568) accounting for a large proportion of the variance increase, in particular for the cohort aged 17 (see Table S3). Not controlling for sex differences yielded a different picture: genetic effects on variance significantly increased between age 11 and age 17 (from .381 to .495; $\Delta\chi^2 = 4.703$; $\Delta df = 1$; $\Delta p = .030$) but did not significantly differ in size between the two older cohorts (.495 versus .526; $\Delta\chi^2 = 0.560$; $\Delta df = 1$; $\Delta p = .454$), consistent with genetic unfolding during adolescence. Controlling for sex differences also reduced twin similarity in Neuroticism (see Figures 3 and S1). These additional analyses indicate that genetic differences in Neuroticism increase in adolescence and unfold as observed trait differences between men and women. Although sex effects also tended to increase for Agreeableness (from .065 to .141) and Conscientiousness (from .062 to .243) across cohorts (with higher scores for females), rerunning the analyses without control for sex differences did not yield a different pattern of results.

Not in line with the expectation that genetic unfolding primarily appears as increasing nonadditive genetic components (hypotheses 3b, 4b, and 5b; see Table 1), we did not find a significant decline in DZ twins' genetic correlations with age, except for Conscientiousness between the youngest cohort and late adolescents. This only significant decline, however, was primarily due to the case that twin similarities indicated shared environmental effects (i.e., $r_{MZ} < 2 \times r_{DZ}$) for the youngest cohort, but not for the older cohorts (cf. Table 4). Thus, our model analyses suggested that the increase in the genetic variance was not attributable

to an enhanced importance of nonadditive genetic contributions, but due to comparable increases in the contributions of both additive and nonadditive genetic sources (see Figure 4 and supplementary Table S4 as well as Figures S2 for more details).

The inconsistency in the results between the previous and the current study may be due to the difference in the modeling strategies. As the nonadditive genetic components were larger than additive genetic components for almost all Big Five traits from the outset in both studies, these would also show larger absolute (but not necessarily relative) increases and these would thus have *a priori* better chances to reach statistical significance. This may explain the significant increases of the nonadditive genetic components based on the variance decomposition model, which Mõttus, Briley and colleagues (in press) relied on. In contrast, the current study focused on variance components corrected for measurement error and tested for the invariance of genetic correlations between DZ twin siblings across cohorts rather than relying on the assumption of either allelic dominance within gene loci or gene-by-gene interactions between gene loci. Thus, the current study did not overestimate any nonadditive in relation to the additive genetic component due to twin similarities attenuated by measurement error or too strict assumptions regarding specific nonadditive genetic sources.

In sum and for the most part, the current study replicates the findings by Mõttus, Briley et al. (in press) revealing significant increases of genetically driven inter-individual trait differences from late childhood to late adolescence and rather constant levels thereafter. This speaks for the relative generalizability of the findings, at least across two Western countries. The findings are in line with the genetic unfolding hypotheses (Table 1) via genetic maturation and/or amplifying contributions of genotype × environment transactions during the transition from childhood to adulthood, when children play an increasingly active role in their own development (Scarr & McCartney, 1983; Kandler & Papendick, 2017). In addition, the increasing genetic variance may also be due to cumulative interaction effects between genotypes and environmental circumstances shared by twins (Purcell, 2002).

Beyond the genetic contribution, our findings suggest specific trends in the magnitude of environmental variation in specific personality traits, such as Extraversion. The SEM analyses yielded significantly lower environmental differences in Extraversion due to environmental sources not shared by twins for the younger cohort (age 11) compared to the two older cohorts (ages 17 and 23). It might be that the unfolding of inter-individual

differences in Extraversion is also due to individualizing differences in experiences not linked with genetic dispositions during adolescence, a crucial time for individuals to find ~~and go~~ their own path in life, detached from parents and siblings.

The model analyses also yielded significantly larger environmental variance in Conscientiousness for the two younger adolescent cohorts compared to the young adult cohort. The decline of environmental differences in Conscientiousness for young adults could reflect the importance of normative environmental pressure to be more conscientious during this period of life, a time of setting the agenda of work life (cf. Denissen et al., 2013). Even though not explicitly tested in their study, Kandler et al. (2010) found a similar decline of true-score variance in Conscientiousness primarily due to environmental trait change and reduced occasional specificity due to individual environmental sources at later measurement occasions, pointing to the robustness of this finding. Significant trends in environmental trait variance are not in line with the only contribution of genetic maturation to the unfolding of personality differences, as proposed by the Five-Factor Theory (FFT) of personality (McCrae & Costa, 2008). In general, however, there was no systematic age trend for the magnitude of environmental variance in Big Five traits, in line with Mõttus, Briley et al. (in press).

Although Study 1 replicated previous findings, the findings are based on a very brief self-report measure. Even though acquiescence was controlled, single short measures and rater perspectives may be subject to further specific method and rater biases distorting the findings (Kandler, 2012b; McCrae, 2015). Also, the cross-sectional design of Study 1 prevents us from allowing more direct inferences about change in the magnitude of inter-individual trait differences and the underlying developmental sources.

Study 2: A Longitudinal Multi-Rater Twin Study on the Genetic and Environmental Sources of the Development of Personality Unfolding from Adolescence to Emerging Adulthood

Longitudinal data from twins are able to discriminate between different accounts for the unfolding of genetic trait variance (see Table 1): Is the existing genetic variance amplified over time or do new genetic sources add to existing genetic differences and thereby increase inter-individual differences? Briley and Tucker-Drob (2013) called these scenarios *amplification* and *innovation*. Amplification of genetic variance over time (across ages) could happen via genotype × environment transactions, whereby genetically driven experiences reinforce genetic differences underlying trait variance. In case of innovation, novel genetic sources of variance may emerge due to activation of gene variants not yet expressed (i.e.,

genetic maturation) and (or) interactions between genotypes and novel environmental circumstances shared by genetically related twins not yet experienced before (e.g., transition from primary to secondary school, more freedom and autonomy from parents for adolescents compared to children and resulting shared experiences, or shared social experiences related to puberty).

Briley and Tucker-Drob (2013) identified amplification as the primary source of increasing genetic differences in intelligence between the ages 8 and 18. However, as genetic variance in intelligence is more stable after childhood (Tucker-Drob & Briley, 2014) than the Big Five genetic variance (Briley & Tucker-Drob, 2014), these findings on intelligence are not necessarily generalizable to personality traits. We addressed this in a longitudinal study of adolescent twins and their transition to adulthood, testing whether amplified genetic differences (i.e., hypothesis 4 in Table 1) or an accumulation of novel genetic factors (i.e., hypotheses 3 and 5 in Table 1) can account for the increase in genetic differences during adolescence.

Since previous studies suggest lower but increasing stability of genetic differences in personality traits during adolescence compared to almost perfect stability of genetic differences in adulthood (Briley & Tucker-Drob, 2017; Kandler & Papendick, 2017) but partly inconsistent results regarding the trends of genetic variance and contributions to trait stability in adolescence (compare Briley & Tucker-Drob, 2014, with Kandler & Papendick, 2017, and Mõttus, Briley et al., in press), we expected support for both genotype × environment transactions and interactions (hypotheses 4 and 5 in Table 1) underlying the increase of genetic variance in personality traits in this period of life. Supplementing twins' self-ratings with mothers' and fathers' ratings allowed us to exclude alternative explanations due to variance in rater biases and rater-specific perspectives (Bartels, Boomsma, Hudziak, van Beijsterveldt, & van den Oord, 2007; Kandler et al., 2010; Zapko-Willmes, Riemann, & Kandler, 2018).

Method

Participants

This study used data from the Norwegian Twin Study of Youths (NTSY)[12], which is a longitudinal twin study on 7 twin birth cohorts born between 1988 and 1994 made available

[12] The NTSY data set is not public domain. It includes a number of resilience-related and psychological health variables. Requests for the data set as scientific use file for own research projects should be sent to Svenn Torgersen and/or Trine Waaktaar, University of Oslo, Norway.

by the National Birth Registry of Norway. The NTSY aimed at a deeper understanding of variation in resilience. From about 5,000 contacted twin families, over 1,500 families provided a signed agreement and *N* = 1,393 participated in the first wave. Questionnaires were sent to the twin families at three points of time two years apart, starting when twins were 12 to 18 years old in the second half of 2006 and the first half of 2007. The second and third waves included the same questionnaires and data were collected when the twins were 14 to 21 and 16 to 23 years old in the second half of 2008 and the first half of 2009 as well as in 2011. From those who participated at the second measurement occasion, *n* = 960 families already provided data at the first point of time and *n* = 105 families did not participate before. In the third wave, data from *n* = 883 families were available including those who participated in all three waves (*n* = 694), only in the first wave (*n* = 102) or in the second wave (*n* = 45), and for the first time (*n* = 42; see Table 5). Twins provided self-reports and parents of twins were asked to fill out the third-person version. That is, parents did not provide data about themselves (except a few demographic questions) but informant reports on each of their twin children (see also Waaktaar, Kan, & Torgersen, 2018).

Measures

Big Five personality traits were measured by means of 40 items from the originally 144 items long Hierarchical Personality Inventory for Children (HiPIC; Mervielde & De Fruyt, 1999). The translation to Norwegian was directed by Vollrath (Norwegian Institute of Public Health). Twins were asked to rate themselves on a 5-point Likert scale (from 0 “not typical of me” to 4 “very typical of me”). Mothers and fathers of twins rated their children on the same items based on the third-person version. The selected items covered the Big Five personality trait dimensions with eight items each. Inter-item consistency ranged between α = .70 and α = .89 across raters and measurement occasions. They tended to be larger for parent reports (αs: .77 - .89) compared to twins’ self-reports (αs: .70 - .86) with marginal variation across measurement occasions (see also supplementary Table S5). Inter-rater correlations tended to be larger among parent reports, ranging from *r* = .49 to *r* = .74, than between parent reports and twins’ self-reports, ranging from *r* = .33 to *r* = .61. Inter-rater consistency ranged between *ICC*(3,3) = .67 and *ICC*(3,3) = .83 (see also supplementary Table S6).

Because of the substantial age range within one measurement occasion and because of potential sex effects that may attenuate opposite-sex DZ twin correlations, we corrected twins’ self-reports as well as parent reports for twins’ age and sex differences within each

point of time using a regression procedure (McGue & Bouchard, 1984). Unstandardized residual scores of self-reports as well as mother and father reports derived from these regressions were used as three reliable indicators of latent traits in the following analyses. DZ twin correlations based on these corrected scores did not differ markedly and systematically between same-sex (*r*: -.02 - .55) and opposite-sex twins (*r*: .01 - .40) across traits, raters, and measurement occasions.

Analyses

All preparatory analyses were run with the statistical software package IBM SPSS 21.0. Similar to Study 1, we used measurement models (see Figure 5A for more details on the measurement level) to disentangle true score (or latent trait τ) variance from random error (ε) variance. True score variance could be estimated as the common variance in twins' self-reports and mothers' as well as fathers' informant reports assuming self-reports and parent reports being τ-congeneric. We also modelled nonrandom variance due to systematic informant-specific perspectives within measurement occasions (i.e., residuals of parent reports are allowed to correlate: m_{MF}) and between points in time (i.e., stable differences due to parent report method factors; cf. Figure 5A). The latter allowed a correction for artificial within-rater biases, such as assimilation or contrast effects due to the fact that the same rater knew and assessed both twin siblings. The modelling of different raters' reliable Big Five trait scores allowed us to more accurately capture personality trait score variance.

The latent trait score variances and residual variances as well as the parent-report residual correlations were allowed to vary across the three measurement occasions. Thus, the SEM enabled tests for metric and strict MI across time. The assumption of metric MI required that factor loadings are invariant across all three measurement occasions (i.e., $\lambda_{M1} = \lambda_{M2} = \lambda_{M3}$, $\lambda_{S1} = \lambda_{S2} = \lambda_{S3}$, and $\lambda_{F1} = \lambda_{F2} = \lambda_{F3}$). The assumption of strict MI, however, required that factor loadings and rater-specific variances, as well as parent-report residual correlations within measurement occasions were invariant across time (i.e., metric MI plus $VAR(\varepsilon_{M1}) = VAR(\varepsilon_{M2}) = VAR(\varepsilon_{M3})$, $VAR(\varepsilon_{S1}) = VAR(\varepsilon_{S2}) = VAR(\varepsilon_{S3})$, and $VAR(\varepsilon_{F1}) = VAR(\varepsilon_{F2}) = VAR(\varepsilon_{F3})$, as well as $m_{MF1} = m_{MF2} = m_{MF3}$; cf. Figure 5).

The used modelling strategy allowed for more accurate biometric model estimates of genetic and environmental sources of latent personality trait differences at each measurement occasion (see Figure 5B for more details on the biometric structure level). We tested different nested models against each other. If novel genetic effects could be fixed to

zero ($g_2 = g_{32} = g_3 = 0$) without significant reduction in model fit and initial genetic effects increased over time ($g_1 < g_{21} < g_{31}$), then this would suggest amplification of initial genetic differences due to genotype × environment transactions (hypothesis 4, see Table 1). The contrary result ($g_2 \neq 0$, $g_{32} \neq 0$, and $g_3 \neq 0$) would support that novel genetic influences occur and accumulate over time, indicating unfolding via genetic maturation (hypothesis 3) that can also depend on changing environmental circumstances shared by twins (hypothesis 5). All genetic unfolding hypotheses can account for increasing genetic variance, if initial genetic factors act constantly or even increase over time ($g_1 \leq g_{21} \leq g_{31}$ and $g_2 \leq g_{32}$) in the presence of novel genetic influences, whereas declining initial genetic variance over time ($g_1 > g_{21} > g_{31}$) in the presence of novel genetic influences would only support the genotype × environment interaction hypothesis (hypothesis 5).

Similarly, we tested whether novel environmental influences contributed to the trait variance (i.e., $e_2 = e_{32} = e_3 = 0$) and whether initial and activated environmental factors remained constant (i.e., $e_1 = e_{21} = e_{31}$ and $e_2 = e_{32}$) over time without specific expectations. If initial environmental factors increased over time (i.e., $e_1 \leq e_{21} \leq e_{31}$) in the presence of novel environmental influences (i.e., $e_2 \neq 0$, $e_{32} \neq 0$, and $e_3 \neq 0$), this would support the environmental individualization hypotheses (see Table 1). Declining environmental components, however, would indicate environmental normalization via social maturation.

As was done in Study 1, the SEM analyses were conducted with the software IBM SPSS AMOS 21.0 (Arbuckle, 2012). Corresponding AMOS scripts are retrievable from the Open Science Framework (https://osf.io/wk9rp/; see folder Study 2). Missing parent reports and dropout over time could be treated as completely-at-random for Extraversion ($\chi^2 = 2474.44$, $df = 2562$, $p = .89$), Openness ($\chi^2 = 2363.24$, $df = 2374$, $p = .56$), and Conscientiousness ($\chi^2 = 2528.09$, $df = 2430$, $p = .08$), but not for Neuroticism ($\chi^2 = 2731.35$, $df = 2559$, $p = .01$) and Agreeableness ($\chi^2 = 2921.28$, $df = 2673$, $p < .001$) based on Missing-Completely-At-Random-tests (Little, 1988). Specific comparisons yielded that fathers tended to provide lower ratings on twins' Neuroticism and Agreeableness. However, dropout over time could be treated as randomly distributed for specific raters. Missing values were handled by full information maximization likelihood (FIML) model fitting procedures (Little & Rubin, 2002).

Results and Discussion

Initial (unconstrained) models provided excellent model fit (RMSEAs < .02 and CFIs > .97). The criteria for measurement model comparisons (i.e., overlapping 90% RMSEA CIs and

ΔCFI < .01) supported strict MI across measurement occasions for all Big Five personality traits (see Table 6). All following hypothesis tests are based on measurement models assuming strict MI.

Initially, we constrained all DZ twins' genetic correlations to be equal across measurement occasions ($\sigma_1 = \sigma_2 = \sigma_3$; cf. Figure 5). For all five traits, the genetic correlation between DZ twin siblings was lower than 0.5 indicating significant nonadditive genetic contributions to the variance. The FIML-based $\Delta\chi^2$-test (comparable to the ML-based $\Delta\chi^2$-test) yielded non-significant reductions of model fit for Neuroticism, Extraversion, Openness, and Agreeableness (values of $\Delta\chi^2 < 5.00$, $\Delta df = 2$, $p > .05$), indicating that the change in the magnitude of genetic variance is not specifically attributable to neither additive nor nonadditive genetic sources. This further contradicts the hypothesis that the nonadditive genetic component rather than the additive genetic component increases from early to late adolescence. For Conscientiousness, the DZ genetic correlation varies significantly across measurement occasions ($\Delta\chi^2 = 6.14$, $\Delta df = 2$, $p = .046$), due to significant differences between times 1 and 2 ($\Delta\chi^2 = 6.14$, $\Delta df = 1$, $p = .013$), but not between times 2 and 3 ($\Delta\chi^2 = 1.77$, $\Delta df = 1$, $p = .183$). The genetic correlations declined from $r_{DZ} = .32$ at time 1 to $r_{DZ} = .05$ at time 2. This result replicates the finding of Study 1, in which we found a similar trend for Conscientiousness between early and late adolescence.

Based on the model assuming strict MI with constrained DZ twins' genetic correlations across measurement occasions (except for Conscientiousness), latent biometric model parameter estimates for all Big Five traits are shown in Table 7 (see also Table S7 for all model parameter estimates including error variances, factor loadings, method factor variances, and all correlations). Constraining novel genetic effects to zero led to a significant decline in model fit for all personality traits. In addition, constraining initial and activated genetic effects to be constant over time also led to a significant decline in model fit for all traits, except Neuroticism. Although genetic differences are largely stable across measurement occasions (ranging between .78 and .95) and account for the most part of the rank-order stability (ranging between 59% and 85%, see Table 8), genetic variance due to initial genetic effects at Time 1 tended to decline over time, on average from 0.17 to 0.14 (see Figure 6). Declining initial genetic differences in the presence of novel genetic sources of variance resulted in balanced heritability estimates: On average .71, .72, and .70 across the three measurement occasions (c.f. Table 8). These findings provide strong support for the

hypotheses 3 and/or 5 and contradict hypothesis 4 (see Table 1) regarding the increase of genetic variance in personality traits. In other words, changes in the magnitude of genetic variance could primarily be due to novel genetic contributions that accumulate over time at least during the age period investigated here (on average age 15 to 20 years).

These model fitting results do not point to the primary role of genotype × environment transactions as a driving force underlying the unfoldment of individual differences. Instead, they point to other developmental mechanisms: Differential genetic maturation (i.e., differential activation of gene variants not yet expressed previously) and interactions between genotypes and novel environmental circumstances not yet experienced before but shared by genetically related twins, because those genotype × environment interactions act like genetic influences (Kandler & Zapko-Willmes, 2017; Briley et al., 2018). Since initial and activated genetic effects tended to decline over time ($g_1 > g_{31}$ and $g_2 > g_{32}$), genetic unfolding via genetic maturation as an exclusive explanation may be less plausible, because it would not predict declines in initial genetic variance, but rather novel genetic factors that add to existing ones. Cumulative genotype × environment interaction effects, however, can account for genetic effects, which both occur at a specific point of time and decline over time. If the unfolding of genetic differences depends on environmental circumstances that can change over time, novel environmental opportunities or limits can change the expression of genetic differences. In other words, MZ twins show a more similar genetic sensitivity to environmental demands (e.g., transition from school to university/training/work life or leaving the parental home) and thus may develop more similarly than DZ twins do. However, changing environmental contexts may not be lasting and new life transitions occur. Thus, interactions between genotypes and environmental circumstances shared by twins, that appear as estimates of genetic contributions to individual differences, may diminish over time and new interaction effects can occur with new environmental opportunities.

In contrast to the developmental pattern of the genetic component, the model analyses did not yield significant novel environmental contributions for Openness and Agreeableness at time 2 and for neither Big Five trait at time 3. The latter resulted in an almost perfect stability of environmental differences between time 2 and 3 (see Table 8). Given the important life stage between age 15 and 20 that includes important life transitions (leaving the parental home, graduation from school, or starting a professional training), which should come along with new social demands and individual events, this finding is

surprising. Since estimates of the environmental contribution to stable differences are relatively small (see Table 8), the findings suggest that environmental influences depend on the genetic sensitivity to them. In other words, genetic and environmental sources are interwoven. In addition, variance due to novel nonshared environmental effects at time 2 tended to amplify from average age 17 to 20 (on average from 0.02 to 0.05 across traits).

More specifically, the analyses yielded different results for different traits, such as increasing environmental variance for Extraversion and declining environmental variance in Conscientiousness. This directly replicates the results of Study 1. In both cases, this was due to novel environmental factors supporting the idea of individualizing influences on Extraversion and normative variance-reducing environmental pressure for Conscientiousness in late adolescents. The current study also provided evidence for environmental increases in Neuroticism variance and declining environmental components for Agreeableness beyond age 17 (cf. Figure 6). In other words, substantial influences that occur before age 17 acting to increase or decrease differences between genetically identical individuals become more conspicuous until the age of 20. It is not clear which individual experiences drive these differences. In any case, however, the findings suggest environmental individualization via amplifying effects of chance for Neuroticism and Extraversion, and environmental pressure that acts to reduce variance in Agreeableness and Conscientiousness in emerging adulthood (see Table 1 and General Discussion for a more in-depth consideration).

In sum, Study 2 replicates most of the patterns derived from Study 1 regarding the increase of trait variance at least up to the average age of 17 across all traits (see Figure 7). This rise came along with an increase in the genetic variance component, whereby neither the additive nor the nonadditive component played an enhanced role. This speaks for the relative generalizability of the findings across two European countries and different study designs. Using a longitudinal design, Study 2 went beyond previous studies in showing that genetic variance increased due to estimates of novel genetic components that add to initial genetic differences in line with the hypotheses of genetic unfolding via genetic maturation or cumulative interaction effects between genotypes and environments shared by siblings raised together, at least during adolescence (see General Discussion for potential examples).

General Discussion

The current investigation provided further evidence for the unfolding of personality differences in youth across two studies using population-based samples from different

European countries and different personality measures, rater perspectives, and study designs. This underscores the replicability and robustness of this phenotypic developmental pattern. The findings further replicate those of Mõttus, Briley and colleagues (in press) who reported a primary role of increasing genetic differences underlying the growing amplitude of personality differences. Our longitudinal Study 2 expands past research providing evidence for novel genetic components that add to existing genetic differences instead of amplification of initial genetic differences via person-environment transactions (or because of nature gradually shining through other influences). In other words, the increase of genetic differences is not simply attributable to an increasingly active role of an individual who expresses his or her genotype in personality-matching environments that shape the individual's trait development (Scarr & McCartney, 1983). Even in the second half of the adolescent period, between the ages of 15 and 20, new genetic components add to the existing genetic differences. This informs about potential genetically influenced mechanisms accounting for personality development in this period of life.

Differential Genetic Maturation

In line with the genetic maturation hypothesis, differential activation of new gene variants not yet expressed previously (e.g., during puberty) may enhance genetic differences underlying observable increases in trait differences (e.g., differential pubertal development). These gene variants may unfold their effects additively but may also interact with other initial and novel genetic factors. As a consequence, both additive and nonadditive genetic variance components gradually increase over time. It may be possible that complex affective, cognitive, and behavioral traits take a longer period of time in contrast to physical features, such as height, to reach their full genetically anchored maturity. Individuals mature at different rates and this may account for the robust finding of differential individual development (rank-order change) due to genetic sources during the first third of life until genetic differences reach almost perfect stability in young adulthood (Briley & Tucker-Drob, 2014; Kandler & Papendick, 2017).

Differential genetic maturation as explanation for novel genetic influences supports popular theories, such as the FFT of personality, which proposes genetic maturation as primary developmental mechanism of personality trait development (McCrae et al., 2000; McCrae & Costa, 2008). However, genetic maturation cannot explain that initial genetic differences tended to gradually diminish across measurement occasions (c.f. Figure 6).

Likewise, if development was merely nature increasingly running its course, as proposed by the FFT, we would have observed no change in the amplitude of the environmental variance. Moreover, development does not take place within an environmental vacuum. Even the development of highly heritable traits, such as height or BMI, depends on environmental opportunities to reach their fully genetically anchored expression (Johnson, 2010). That is, genetic unfolding depends on the opportunities and limits of the environmental conditions, leaving hypothesis 5 in Table 1 as most plausible explanation of increasing genetic variance.

Genotype × Environment Interaction Effects

The range of available environments determines which opportunities and limits individuals have to express themselves, to find or select their optimal niche, or to avoid inappropriate social and economic conditions (Scarr, 1992, 1993). Adolescent twin siblings raised in the same family are likely to effectively share experiences that vary between families (e.g., resulting from common peers, shared clubs, same schooling and living conditions). Interactions between these shared experiences and heritable tendencies would act to increase within-family similarity as a function of the genetic resemblance, because different genetic dispositions may be differently sensitive to same environmental conditions (Belsky & Pluess, 2009; Kandler & Ostendorf, 2016; Krueger et al., 2008). That is, MZ twins would become more similar relative to DZ twins. Since new opportunities and chances that occur during development and socio-biological life transitions (e.g., pubertal change) can alter the genetic sensitivity to environmental stressors, genotype × environment interaction effects may accumulate to some degree over time, but need not necessarily last on the long run. The latter accounts for the consistent finding that initial genetic differences tended to gradually decline over time. *Cumulative* and *altering genotype × environment interaction effects* shared by twins can account for both estimates of novel genetic sources that act to increase genetic differences in adolescent personality traits and nonstable genetic differences during this life period.

Although Study 2 did not find support for the hypothesis that initial genetic differences amplify over time due to genotype × environment transactions, few would deny that people become more and more autonomous and self-directed during adolescence. They have increasing opportunities to pave their own way, evoke, select, and create environments that match their heritable tendencies. These environments in turn can provide experiences that have the potential to reinforce the pre-existing tendencies and stabilize inter-individual trait

differences (McAdams, 2015). However, not every person may have the same chance to unfold his or her genetic makeup depending on environmental circumstances. Similarly, genetic predispositions to extreme personality traits may be more or less sensitive to specific environmental stimuli than trait levels closer to a normative population level (Mõttus, Briley et al., in press). For example, people high on neuroticism may perceive the world as a more negative place and experience (or report) more negative life events (i.e. *transaction*), but the probability to develop depression does not only increase with the level of neuroticism and stressful life events, but also depend on the availability of protective positive experiences (i.e. *interaction*; see Kandler & Ostendorf, 2016). According to these considerations, cumulative genotype × environment interaction effects may rather reflect *differential genotype* × *environment transactions* × *environment interactions*.

Environmental opportunities may also change over time. When late adolescents or young adults leave their parental home and go their own ways, the probability of environmental conditions shared by twins declines, whereas the probability of nonshared contextual circumstances increases. Thus, the contribution of potential interaction effects between genotypes and environmental conditions shared by twins would decrease and the contribution of interaction effects between genotypes and nonshared environmental conditions would increase over time. Since genotype × environment interaction effects shared by twins (if not directly estimated in twin studies) would appear as genetic component, whereas interaction effects not shared by twins would be confounded with estimates of nonshared environmental effects (Purcell, 2002), the shift from shared to nonshared interaction effects in young adulthood can explain previous findings on reducing genetic differences and increasing variance due to environmental individualizing in adulthood (Briley & Tucker-Drob, 2014; Kandler & Papendick, 2017).

Genotype (× environment transaction) × environment interaction effects may create a balance between declining genetic variance and increasing individualizing environmental effects that can explain the constant level of inter-individual differences in adult personality traits (Mõttus et al., 2016). In sum, genotype × environment interaction effects that can accumulate but also change to some degree across age represent a promising explanation for (1) estimates of novel genetic influences in late adolescence, (2) the increase of genetic differences from childhood to adolescence, (3) the plateauing and stabilization of genetic variance in emerging adulthood, and (4) reducing heritability across the adult age. Future

studies should test this hypothesis with a broader age range and in different life stages, as different mechanisms may operate differently in different phases of life or as Mõttus et al. (2017) put it: "the developmental mechanisms underlying increases in youth personality variance may operate more strongly during childhood than adolescence" (p. 323).

Measurement Artifact

An alternative explanation for increasing variance in personality trait scores and associated genetic variance may be an increase in the reliability and accuracy of personality measurement. Even though we could establish strict measurement invariance for HiPIC trait scores based on three rater perspectives across three measurement occasions between the average ages of 15 and 20 in Study 2, our model analyses did not allow to assume strict MI for BFI-S trait scores across birth cohorts with the average ages of 11, 17, and 23 in Study 1. Previous studies have reported profound differences between young and older adolescents in the accuracy of capturing Big Five traits (e.g., Allik, Laidra, Realo, & Pullmann, 2004; Mõttus et al., 2015; Tackett et al., 2012) and that heritability estimates increase with the accuracy of measurement (e.g., Kandler, 2012b; Kandler & Papendick, 2017).

Possibly both measurement artifact and typical developmental processes may be interwoven accounting for the emergence of inter-individual personality differences between childhood and adolescence. In the second decade of life, adolescents may have enriched and more accurate self-views as well as more opportunities and the capacity to express their personality traits. This may allow them to provide more accurate self-reports and well-acquainted others to provide more accurate informant reports on their personality. This produces artifactual differences between self-reports from children and those from adolescents. However, it is less plausible to explain age trends based on parent reports with those processes. Since Mõttus, Briley et al. (in press) found comparable age differences in the amplitude of genetic variance across self- and parent reports, the sheer artifact account may not be the exclusive explanation.

Amplifying Environmental Individualization

Neither measurement invariance nor cumulative genotype × environment interaction effects not shared by twins, however, can account for there being no novel nonshared environmental influences at the third measurement occasion (average age 20) in Study 2 for none of the traits, whereas novel environmental influences at time 2 (average age 17) tended to increase between time 2 and 3 (between ages 17 and 20) for all traits (except

Openness). This is consistent with there being individualizing influences that occur intra-individually by chance at some time and can generate growth processes, which amplify inter-individual differences over time (Molenaar et al., 1993, 2003). One random event may affect the probability of further experiences in a very idiosyncratic way depending on the very individual situation of people; Mõttus, Briley et al. (in press) called this the "random walk" hypothesis. For example, a lottery win may come along with stronger changes of circumstances for those people with lower socioeconomic status compared to those with a higher status. Both social backgrounds may produce different effects. Even though genetically identical twins share the same interest (e.g., to go to university and study the same subject) chance may affect differences in how they can realize their interests (e.g., different universities in different cities under different social, economic, and ecological conditions) resulting in highly idiosyncratic life journeys with unique experiences affecting individualization.

Random life events could have small *normative* effects in the short run, but strong *idiographic* effects on the long run via amplifying effects of several consequences that result from an initial event. Those idiographic chains of causation may primarily act as nonshared environmental influences that increase over time. They can even manifest epigenetically generating variation in neural network structures and functioning, and thus in affective, cognitive, motivational, behavioral differences (Gottlieb, 2003; Kandler & Zapko-Willmes, 2017). Environmental sources can switch on and off the genetic activity without altering the genome, but shaping the epigenome referred to as epigenetic regulation. About one third of stable epigenetic differences and dynamic repeatability of epigenetic regulation over the human life course have been found to be modifiable by the environment (Shah et al., 2014). Increasing epigenetic differences between genetically identical MZ twins – referred to as "epigenetic drift" – that arise during development have been found to be primarily driven by individually unique (nonshared) environmental influences (Tan et al., 2016). Our findings are consistent with the idea that novel epigenetic differences in personality traits occur during puberty and amplify over time, a speculation that needs to be addressed by future studies.

Different random events, however, may not act as systematic accumulation of isolated chances, but as a chaotic process of continuously changing random constellations and interactions of individualizing sources that appear as consistently novel unsystematic influences. This should rather result in a random variation around a specific set point and act

to decrease inter-individual differences over time (see Mõttus et al., 2017, for a simulation study). Those sources of variance are hard to capture, because one event may change the probability of further experiences in a very idiosyncratic and chaotic way depending on the very individual situation of people (Kandler, Zapko-Willmes, Richter, & Riemann, in press; Luhmann, Orth, Specht, Kandler, & Lucas, 2014).

Normative Environmental Pressure

Neither the "random walk" nor the "epigenetic drift" explanation for environmental individualization can account for the robustly replicable environmentally driven decline of inter-individual differences in Conscientiousness (and Agreeableness in Study 2) after reaching the adult age. This points to the possibility that normative environmental pressure can act to increase Conscientiousness on average as well as reduce its variation. People do not only seek to reach their own inherent values and goals, but also follow social values and standards that set the direction of personality development (Denissen et al., 2013). In emerging adulthood, when young persons have to gain important educational and occupational qualifications, they have to show a certain level of orderliness and discipline as well as a certain sense of duty and responsibility to create the social and economic basis for what most people want in their life: An own family, social participation, and a satisfying job.

Although our analyses did not support the role of normative social pressure for all Big Five personality traits, it may act in addition to the other mentioned developmental sources of trait variance and may counterbalance them. This may be particularly the case for those traits that show age trends towards socially functional maturity (Bleidorn et al., 2013) and could account for constant levels of personality trait variation across the adult years of professional activity (Loehlin & Martin, 2001; Mõttus et al., 2016). The role of socialization processes could be tested by investigating age differences in trait variance in a larger sample of traits and testing if decreasing variance is especially likely for traits that are subject to strong socialization pressures (e.g., as rated by experts).

Strengths, Limitations, and Future Directions

The current investigation has several strengths. Two large and population-based twin-family samples from two European countries increased the robustness of our findings. We combined a birth-cohort twin study and a longitudinal twin design to replicate and add to the findings from previous studies. We further used a latent variable approach based on different personality indicators and perspectives taking measurement artifact and invariance

into account. Despite these strengths, both studies have specific limitations that need to be addressed by future studies.

Although Study 2 overcame several limitations of Study 1 by using more comprehensive personality trait measures, multiple informants, and a longitudinal design, the general age range was comparatively small and primarily captured the period after puberty. In Study 1, however, and Mõttus, Briley et al. (in press) highlighted that the most important changes regarding the amplitude of inter-individual differences in personality traits take place before the age of 16. Therefore, further longitudinal studies capturing this important period of personality development are crucial. In addition, genetically informative longitudinal studies spanning different adult age ranges are required to inform about the sources of the amplitude of inter-individual trait differences in other life phases, such as old age (e.g., Allemand, Zimprich, & Martin, 2008; Pedersen & Reynolds, 1998).

Both studies relied on Big Five personality trait measures to capture the personality construct. As already mentioned, measurement issues may partly account for lower trait variance in childhood compared to adulthood. Moreover, several studies have pointed to the validity and usefulness of more specific levels of abstraction within a broad personality system (Mõttus, Sinick et al., in press; Seeboth & Mõttus, 2018). Thus, future studies using more nuanced and comparably valid measures of personality traits for different age groups (i.e., assessments that are fair with respect to age differences) may allow a more detailed and accurate view on the sources of unfolding, stabilization, and reducing trait variation across the lifespan.

Although we controlled for main effects of sex differences on personality variance and found them to be primarily genetic (in case of Neuroticism), we did not focus on sex differences in genetic and environmental sources of personality trait differences. As sex differences arise during pubertal development, a more specific focus on the sources of male and female personality differences could be an interesting endeavor for future studies.

Conclusion

Two studies helped to unravel the complex interplay between genetic and environmental contributions in the unfolding of inter-individual differences in personality traits between early adolescence and young adulthood. We found evidence for the unfolding of personality differences in adolescence, in particular before the age of 17. This increase of inter-individual differences was primarily due to increases in genetic differences. The

increase in genetic variance, in turn, was found to be primarily due to novel genetic sources adding to and partly replacing existing ones. These findings are consistent with complex interaction effects between individual genotypes and environmental circumstances shared by twins that can accumulate over time, but also change (e.g., a shift from shared to nonshared opportunities) and need not last on the long run. Thus, the changing genetic variance component in personality traits over time may reflect a complex genotype × environment interplay. Both studies also provide robust evidence for developmental patterns specific to certain personality traits, such as the decline of environmental variance in Conscientiousness in emerging adulthood, pointing to normative environmental pressure at least to show a minimum of being conscientious during this important life transition. We hope that our investigation paves the way for further studies with broader and different age ranges that examine the specific developmental mechanisms underlying the increase, stabilization, or even decrease of personality variance across the entire lifespan.

References

Allemand, M., Zimprich, D., & Hendriks, A. A. J. (2008). Age differences in five personality domains across the life span. *Developmental Psychology*, *44*, 758-770. https://doi.org/10.1037/0012-1649.44.3.758

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

Allik, J., Laidra, K., Realo, A., & Pullmann, H. (2004). Personality development from 12 to 18 years of age: Changes in mean levels and structure of traits. *European Journal of Personality*, *18*, 445-462. https://doi.org/10.1002/per.524

Arbuckle, J. L. (2012). *IBM SPSS AMOS (Version 21.0) [Computer Program]*. Chicago, IL: IBM.

Arnett, J. J. (2000). Emerging adulthood: A theory of development from late teens through the twenties. *American Psychologist*, *55*, 469-480. https://doi.org/10.1037//0003-066X.55.5.469

Bartels, M., Boomsma, D. I., Hudziak, J. J., van Beijsterveldt, T. C. E. M., & van den Oord, E. J. G. (2007). Twins and the study of rater (dis)agreement. *Psychological Methods*, *12*, 451-466. https://doi.org/10.1037/1082-989X.12.4.451

Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin*, *135*, 885-908. https://doi.org/10.1037/a0017376

Bleidorn, W., Klimstra, T. A., Denissen, J. J. A., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2013). Personality maturation around the world: A cross-cultural examination of social-investment theory. *Psychological Science*, *24*, 2530-2540. https://doi.org/10.1177/0956797613498396

Bratko, D., & Butković, A. (2007). Stability of genetic and environmental effects from adolescence to young adulthood: Results of Croatian longitudinal twin study of personality. *Twin Research and Human Genetics, 10,* 151-157. http://dx.doi.org/10.1375/twin.10.1.151

Briley, D. A., & Livengood, J., & Derringer, J. (2018). Behavior genetic frameworks of causal reasoning for personality psychology. *European Journal of Personality*, *32*, 202-220. https://doi.org/10.1002/per.2153

Briley, D. A., & Tucker-Drob, E. M. (2013). Explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science*, *24*, 1704-1713. http://dx.doi.org/10.1177/0956797613478618

Briley, D. A., & Tucker-Drob, E. M. (2014). Genetic and environmental continuity in personality development: A meta-analysis. *Psychological Bulletin*, *140*, 1303-1331. http://dx.doi.org/10.1037/a0037091

Briley, D. A., & Tucker-Drob, E. M. (2017). Comparing the developmental genetics of cognition and personality over the life span. *Journal of Personality, 85,* 51-64. http://dx.doi.org/10.1111/jopy.12186

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

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233-255. http://doi.org/10.1207/S15328007SEM0902_5

Denissen, J. J. A., Ulferts, H., Lüdtke, O., Muck, P. M., & Gerstorf, D. (2014). Longitudinal transactions between personality and occupational roles: A large and heterogeneous study of job beginners, stayers, and changers. *Developmental Psychology*, *50*, 1931-1942. http://dx.doi.org/10.1037/a0036994

Denissen, J. J. A., Van Aken, M. A. G., Penke, L., & Wood, D. (2013). Self-regulation underlies temperament and personality: An integrative developmental framework. *Child Development Perspectives*, *7*, 255-260. https://doi.org/10.1111/cdep.12050

Eaves, L. J., Last, K., Martin, N. G., & Jinks, J. L. (1977). A progressive approach to non-additivity and genotype-environmental covariance in the analysis of human differences. *British Journal of Mathematical and Statistical Psychology*, *30*, 1-42. https://doi.org/10.1111/j.2044-8317.1977.tb00722.x

Gillespie, N. A., Evans, D. E., Wright, M. M., & Martin, N. G. (2004). Genetic simplex modeling of Eysenck's dimensions of personality in a sample of young Australian twins. *Twin Research, 7,* 637-648. http://dx.doi.org/10.1375/1369052042663814

Gottlieb, G. (2003). On making behavioral genetics truly developmental. *Human Development*, *46*, 337-355. https://doi.org/10.1159/000073306

Hahn, E., Gottschling, J., Bleidorn, W., Kandler, C., Spengler, M., Kornadt, A. E., … Spinath, F. M. (2016). What drives the development of social inequality over the life course? The German TwinLife Study. *Twin Research and Human Genetics*, *19*, 659-672. https://doi.org/10.1017/thg.2016.76

Hahn, E., Gottschling, J., & Spinath, F. M. (2012). Short measurements of personality – Validity and reliability of the GSOEP Big Five Inventory (BFI-S). *Journal of Research in Personality*, *46*, 355-359. http://dx.doi.org/10.1016/j.jrp.2012.03.008

Harden, K. P., Tucker-Drob, E. M., & Tackett, J. L. (2013). The Texas Twin Project. *Twin Research and Human Genetics, 16,* 385-390. https://doi.org/10.1017/thg.2012.97

Hopwood, C. J., Donnellan, M. B., Krueger, R. F., McGue, M., Iacono, W. G., Blonigen, D. M., & Burt, S. A. (2011). Genetic and environmental influences on personality trait stability and growth during the transition to adulthood: A three-wave longitudinal study. *Journal of Personality and Social Psychology, 100*, 545-556. https://doi.org/10.1037/a0022409

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

Jeronimus, B. F., Riese, H., Sanderman, R., & Ormel, J. (2014). Mutual reinforcement between neuroticism and life experiences: A five-wave, 16-year study to test reciprocal causation. *Journal of Personality and Social Psychology*, *107*, 751-764. http://dx.doi.org/10.1037/a0037009

Johnson, A. M., Vernon, P. A., & Feiler, A. R. (2008). Behavioral genetic studies of personality: An introduction and review of the results of 50+ years of research. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The sage handbook of personality theory and assessment* (Personality theories and models, Vol. 1, pp. 145–173). London: Sage.

Johnson, W. (2010). Understanding the genetics of intelligence: Can height help? Can corn oil? *Current Directions in Psychological Science*, *19*, 177-182. http://dx.doi.org/10.1177/0963721410370136

Kandler, C. (2012a). Nature and nurture in personality development: The case of neuroticism and extraversion. *Current Directions in Psychological Science*, *21*, 290-296. http://dx.doi.org/10.1177/0963721412452557

Kandler, C. (2012b). Knowing your personality is knowing its nature: The role of information accuracy of peer assessments for heritability estimates of temperament and personality

traits. *Personality and Individual Differences*, *53*, 387-392. http://dx.doi.org/10.1016/j.paid.2012.01.004

Kandler, C., Bleidorn, W., Riemann, R., Angleitner, A., & Spinath, F. M. (2012). Life events as environmental states and genetic traits and the role of personality: A longitudinal twin study. *Behavior Genetics*, *42*, 57-72. http://dx.doi.org/10.1007/s10519-011-9491-0

Kandler, C., Bleidorn, W., Riemann, R., Spinath, F. M., Thiel, W. & Angleitner, A. (2010). Sources of cumulative continuity in personality: A longitudinal multiple-rater twin study. *Journal of Personality and Social Psychology*, *98*, 995-1008. http://dx.doi.org/10.1037/a0019558

Kandler, C., & Ostendorf, F. (2016). Additive and synergetic contributions of neuroticism and life events to depression and anxiety in women. *European Journal of Personality*, *30*, 390-405. http://dx.doi.org/10.1002/per.2065

Kandler, C., & Papendick, M. (2017). Behavior genetics and personality development: A methodological and meta-analytic review. In: J. Specht (ed.), *Personality Development across the Lifespan* (pp. 473-495). Elsevier Academic Press. http://dx.doi.org/10.1016/B978-0-12-804674-6.00029-6

Kandler, C., & Zapko-Willmes, A. (2017). Theoretical perspectives on the interplay of nature and nurture in personality development. In: J. Specht (ed.), *Personality Development across the Lifespan* (pp. 101-115). Elsevier Academic Press. http://dx.doi.org/10.1016/B978-0-12-804674-6.00008-9

Kandler, C., Zapko-Willmes, A., Richter, J., & Riemann, R. (in press). Synergistic and dynamic interdependence within and between genetic and environmental factors of personality differences. In: J. F. Rauthmann (ed.), *The Handbook of Personality Dynamics and Processes*. Elsevier Academic Press.

Krueger, R. F., South, S., Johnson, W., & Iacono, W. G. (2008). The heritability of personality is not always 50%: Gene-environment interactions and correlations between personality and parenting. *Journal of Personality*, *76*, 1485-1522. http://dx.doi.org/10.1111/j.1467-6494.2008.00529.x

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of American Statistical Association, 83,* 1198-1202. http://dx.doi.org/10.2307/2290157

Little, R. J. A., & Rubin, D. B. (2002). Statistical analysis with missing data (2nd ed.). Hoboken, NJ: Wiley.

Loehlin, J. C., & Martin, N. G. (2001). Age changes in personality traits and their heritabilities during the adult years: Evidence from Australian Twin Registry samples. *Personality and Individual Differences, 30*, 1147-1160. http://dx.doi.org/10.1016/S0191-8869(00)00099-4

Lüdtke, O., Roberts, B. W., Trautwein, U., & Nagy, G. (2011). A random walk down university avenue: Life paths, life events, and personality trait change at the transition to university life. *Journal of Personality and Social Psychology*, *101*, 620-637. http://dx.doi.org/10.1037/a0023743

Luhmann, M., Orth, U., Specht, J., Kandler, C., & Lucas, R. E. (2014). Studying changes in life circumstances and personality: It's about time. *European Journal of Personality*, *28*, 256-266. http://dx.doi.org/10.1002/per.1951

Lykken, D. T. (1982) Research with twins: the concept of emergenesis. *Psychophysiology*, 19, 361-373. https://doi.org/10.1111/j.1469-8986.1982.tb02489.x

Lykken, D. T. (2006). The mechanism of emergenesis. *Genes, Brain, and Behavior*, *5*, 306-310. https://doi.org/10.1111/j.1601-183X.2006.00233.x

MacCallum, R., Browne, M., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *2*, 130-149. http://dx.doi.org/10.1037/1082-989X.1.2.130

McAdams, D. P. (2015). *The art and science of personality development*. New York: Guilford.

McAdams, D. P., & Olson, B. D. (2010). Personality development: Continuity and change over the life course. *Annual Review of Psychology*, *61*, 517-542. http://dx.doi.org/10.1146/annurev.psych.093008.100507

McCrae, R. R. (2015). A more nuanced view of reliability: specificity in the trait hierarchy. *Personality and Social Psychology Review*, *19*, 97-112. http://dx.doi.org/10.1177/1088868314541857

McCrae, R. R., & Costa, P. T. (2008). The Five-Factor Theory of personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 159_181). New York, NY: Guilford Press.

McCrae, R. R., Costa, P. T., de Lima, M. P., Simões, A., Ostendorf, F., Angleitner, A., … Piedmont, R. L. (1999). Age differences in personality across the adult life span: Parallels

in five cultures. *Developmental Psychology*, *35*, 466-477. http://dx.doi.org/10.1037/0012-1649.35.2.466

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,* 173-186. http://dx.doi.org/10.1037/0022-3514.78.1.173

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*, 175-215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

McGue, M., & Bouchard, T. J. Jr. (1984). Adjustment of twin data for the effects of age and sex. *Behavior Genetics*, *14*, 325-343. https://doi.org/10.1007/BF01080045

Mervielde, I. & De Fruyt, F. (1999). Construction of the Hierarchical Personality Inventory for Children (HiPIC). In I.Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (pp. 107-127). Tilburg, Netherlands: Tilburg University Press.

Molenaar, P. C. M., Boomsma, D. I., & Dolan, C. V. (1993). A third source of developmental differences. *Behavior Genetics, 23,* 519-524. https://doi.org/10.1007/BF01068142

Molenaar, P.C.M., Huizinga, H.M., & Nesselroade, J.R. (2003). The relationship between the structure of inter-individual and intra-individual variability: A theoretical and empirical vindication of developmental systems theory. In U.M. Staudinger & U. Lindenberger (Eds.), *Understanding human development* (pp. 339.360). Boston: Kluwer.

Mõttus, R. (2017). Five-Factor Theory and personality development. In: J. Specht (ed.), *Personality Development across the Lifespan* (pp. 101-115). Elsevier Academic Press. http://dx.doi.org/10.1016/B978-0-12-804674-6.00007-7

Mõttus, R., & Allerhand, M. (2018). Why do traits come together? The underlying trait and network approaches. In: V. Zeigler-Hill & T. K. Shackelford (eds.), *SAGE handbook of personality and individual differences: Volume 1. The science of personality and individual differences* (pp. 130–151). Los Angeles, CA: Sage.

Mõttus, R., Allerhand, M., & Johnson, W. (2017). Computational modeling of person-situation transactions: How accumulation of situational experiences can shape the distributions of trait scores. In D. C. Funder, J. F. Rauthmann, & R. A. Sherman (Eds.), The *Oxford handbook of psychological situations*. Oxford, England: Oxford University Press.

http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780190263348.001.0001/oxfordhb-9780190263348

Mõttus, R., Allik, J., Hřebíčková, M., Kööts-Ausmees, L., & Realo, A. (2016). Age Differences in the Variance of Personality Characteristics. *European Journal of Personality*, *30*, 4-11. http://dx.doi.org/10.1002/per.2036

Mõttus, R., Briley, D., Zheng, A., Mann, F. D., Engelhardt, L. E., Tackett, J. L., … Tucker-Drob, E. M. (in press). Kids becoming less alike: A behavioral genetic analysis of developmental increases in personality variance from childhood to adolescence. *Journal of Personality and Social Psychology*. http://dx.doi.org/10.1037/pspp0000194

Mõttus, R., Realo, A., Allik, J., Esko, T., Metspalu, A., & Johnson, W. (2015). Within-trait heterogeneity in age group differences in personality domains and facets: Implications for the development and coherence of personality traits. *PLoS ONE*, *10*, e0119667. https://doi.org/10.1371/journal.pone.0119667

Mõttus, R., Sinick, J., Terracciano, A., Hřebíčková, M., Kandler, C., Ando, J., Mortensen, E. L., Colodro-Conde, L., & Jang, K. L. (in press). Personality characteristics below facets: A replication and meta-analysis of cross-rater agreement, rank-order stability, heritability and utility of personality nuances. *Journal of Personality and Social Psychology*. http://dx.doi.org/10.1037/pspp0000202

Mõttus, R., Soto, C. J., & Slobodskaya, H. R. (2017). Are all kids alike? The magnitude of individual differences in personality characteristics tends to increase from early childhood to early adolescence. *European Journal of Personality*, *31*, 313-328. http://dx.doi.org/10.1002/per.2107

Pedersen, N. L., & Reynolds, C. A. (1998). Stability and change in adult personality: Genetic and environmental components. *European Journal of Personality*, *12*, 365-386. https://doi.org/10.1002/(SICI)1099-0984(1998090)12:5<365::AID-PER335>3.0.CO;2-N

Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (2nd ed., Vol. 1, pp. 703-732). New York: Wiley.

Pilia, G., Chen, W.-M., Scuteri, A., Orrú, M., Albei, G., Die, M., … Schlessinger, D. (2006). Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genetics*, *2*, e132. https://doi.org/10.1371/journal.pgen.0020132

Plomin, R., Corley, R., Caspi, A., Fulker, D. W., & DeFries, J. (1998). Adoption results for self-reported personality: Evidence for nonadditive genetic effects? *Journal of Personality and Social Psychology, 75,* 211-218. http://dx.doi.org/10.1037/0022-3514.75.1.211

Plomin, R., & Daniels, D. (2011). Why are children in the same family so different from one another? *International Journal of Epidemiology*, *40*, 563-582. https://doi.org/10.1093/ije/dyq148

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype–environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*, 309-322. https://doi.org/10.1037/0033-2909.84.2.309

Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., . . . Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*, 702-709. http://dx.doi.org/10.1038/ng.3285

Purcell, S. (2002). Variance components models for gene-environment interaction in twin analysis. *Twin Research and Human Genetics*, *5*, 554-571. http://dx.doi.org/10.1375/136905202762342026

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2003). Work experiences and personality development in young adulthood. *Journal of Personality and Social Psychology*, *84*, 582-593. http://dx.doi.org/10.1037/0022-3514.84.3.582

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. http://dx.doi.org/10.1037/0033-2909.126.1.3

Roberts, B. W., & Nickel, L. B. (2017). A critical evaluation of the Neo-Socioanalytic Model of personality. In J. Specht (Ed.), *Personality development across the life-span* (pp. 157-177). Elsevier Academic Press. Elsevier Academic Press. http://dx.doi.org/10.1016/B978-0-12-804674-6.00011-9

Roberts, B. W., & Robins, R. W. (2004). Person-environment fit and its implications for personality development: A longitudinal study. *Journal of Personality*, 72, 89-110. http://dx.doi.org/10.1111/j.0022-3506.2004.00257.x

Roberts, B. W., Walton, K. E., Bogg, T., & Caspi, A. (2006). De-investment in work and non-normative personality trait change in young adulthood. *European Journal of Personality*, *20*, 461-474. http://dx.doi.org/10.1002/per.607

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. http://dx.doi.org/10.1037/0033-2909.132.1.1

Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development*, *63*, 1-19. http://dx.doi.org/10.1111/j.1467-8624.1992.tb03591.x

Scarr, S. (1993). Biological and cultural diversity: The legacy of Darwin for development. *Child Development*, *64*, 1333-1353. http://dx.doi.org/10.2307/1131538

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54,* 424-435. http://dx.doi.org/10.2307/1129703

Seeboth, A., & Mõttus, R. (2018). Successful explanations start with accurate descriptions: Questionnaire items as personality markers for more accurate predictions. *European Journal of Personality*, *32*, 186-201. https://doi.org/10.1002/per.2147

Shah, S., McRae, A. F., Marioni, R. E., Harris, S. E., Gibson, J., Henders, A. K., … Visscher, P. M. (2014). Genetic and environmental exposures constrain epigenetic drift over the human life course. *Genome Research*, *24*, 1725-1733. http://dx.doi.org/10.1101/gr.176933.114

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of Big Five self-reports: acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, *94*, 718-737. http://dx.doi.org/10.1037/0022-3514.94.4.718

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, *100*, 330-348. http://dx.doi.org/10.1037/a0021717

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

Spengler, M., Gottschling, J., & Spinath, F. M. (2012). Personality in childhood: A longitudinal behavior genetic approach. *Personality and Individual Differences, 53*, 411-416. http://dx.doi.org/10.1016/j.paid.2012.01.019

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25*, 173-180. https://doi.org/10.1207/s15327906mbr2502_4

Tackett, J. L., Slobodskaya, H. R., Mar, R. A., Deal, J., Halverson, C. F., Baker, S. R., … Besevegis, E. (2012). The hierarchical structure of childhood personality in five countries: Continuity from early childhood to early adolescence. *Journal of Personality*, *80*, 847-879. https://doi.org/10.1111/j.1467-6494.2011.00748.x

Tan, Q., Heijmans, B. T., Hjelmborg, J. v. B., Soerensen, M., Christensen, K., & Christiansen, L. (2016). Epigenetic drift in the aging genome: a ten-year follow-up in an elderly twin cohort. *International Journal of Epidemiology*, 45, 1146-1158. https://doi.org/10.1093/ije/dyw132

Tanner, J. M., Whitehouse, R. H., & Takaishi, M. (1966). Standards from birth to maturity for height, weight, height velocity, and weight velocity: British children, 1965. *Archives of Disease in Childhood, 41,* 454-471. https://doi.org/10.1136/adc.41.219.454

Tucker-Drob, E. M., & Briley, D. A. (2014). Continuity of genetic and environmental influences on cognition across the life span: A meta-analysis of longitudinal twin and adoption studies. *Psychological Bulletin*, *140*, 949-979. http://dx.doi.org/10.1037/a0035893

Turkheimer, E., & Waldron, M. (2000). Nonshared environment: a theoretical, ethodological, and quantitative review. *Psychological Bulletin*, *126*, 78-108. http://dx.doi.org/10.1037/0033-2909.126.1.78 nonshared

Viken, R. J., Rose, R. J., Kaprio, J., & Koskenvuo, M. (1994). A developmental genetic analysis of adult personality: Extraversion and neuroticism from 18 to 59 years of age. *Journal of Personality and Social Psychology*, *66*, 722-730. http://dx.doi.org/10.1037/0022-3514.66.4.722

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era – concepts and misconceptions. *Nature Review Genetics*, *9*, 255-266. http://dx.doi.org/10.1038/nrg2322

Vukasović, T., & Bratko, D. (2015). Heritability of personality: A meta-analysis of behavior genetic studies. *Psychological Bulletin*, *141*, 769-785. http://dx.doi.org/10.1037/bul0000017

Waaktaar, T., Kan, K.-J., & Torgersen, S. (2018). The genetic and environmental architecture of substance use development from early adolescence into young adulthood: a longitudinal twin study of comorbidity of alcohol, tobacco and illicit drug use. *Addiction*, *113*, 740-748. https://doi.org/10.1111/add.14076

Weiss, A., & King, J. E. (2015). Great ape origins of personality maturation and sex differences: A study of orangutans and chimpanzees. *Journal of Personality and Social Psychology*, *108*, 648-664. http://dx.doi.org/10.1037/pspp0000022

Zapko-Willmes, A., Riemann, R., & Kandler, C. (2018). Unravelling quasi-causal environmental effects via phenotypic and genetically informed multi-rater models: The case of differential parenting and authoritarianism. *European Journal of Personality*, *32*, 233-253. http://dx.doi.org/10.1002/per.2144

Zimmermann, J., & Neyer, F. J. (2013). Do we become a different person when hitting the road? Personality development of sojourners. *Journal of Personality and Social Psychology*, *105*, 515-530. http://dx.doi.org/10.1037/a0033019

Table 1. *Overview of Hypotheses Accounting for Trends in Trait Variance over Time*

Hypotheses	Explanations and consequences for estimates of genetic and environmental trait variance and rank-order stability
(1) *Environmental individualization* via amplifying effects of chance	Initial environmental differences increase over time resulting in amplified environmental contributions to trait stability
(2) *Environmental individualization* via accumulation of life experiences or interactions between life events	Environmental differences increase via novel experiences that add to or interact with existing ones resulting in reduced environmental stability
(3a/b) *Genetic unfolding* via genetic maturation	Innate tendencies unfold and stabilize with development, novel genetic effects will be activated and accumulate during development
(4a/b) *Genetic unfolding* via amplifying effects of genotype × environment transactions	Initial genetic differences expand and stabilize over time via person × environment transactions and thus via dynamic interplays between genetic differences and nonrandom environmental circumstances resulting in amplified genetic contributions to trait stability
(5a/b) Accumulation of interaction effects between genotypes and environments shared by twins/siblings (appears as *genetic unfolding*)	Estimates of genetic variance increase (decrease) with the opportunities (limits) to express associated tendencies (individuals develop more/less similar as a function of their genetic relatedness), appear as novel genetic effects that accumulate over time
(6) Accumulation of interaction effects between genotypes and environments not shared by twins/siblings (appears as *environmental individualization*)	Estimates of environmental variance increase (decrease) with the opportunities (limits) to express associated tendencies (individuals become less similar irrespective of their genetic relatedness), appear as novel environmental effects not shared by twins that accumulate over time
(7) *Environmental normalization* via social/cultural maturation	Normative environmental pressure act to reduce trait differences, appears as a decline of the environmental component

Note. Genetic unfolding can appear as increasing additive (a) or nonadditive (b) genetic variance, or both. Genotype × environment transactions would not act to increase genetic variance per se, but the expression of genetic variation in fitting environments, resulting in an enhanced probability of environmental influences aligned with genetic differences.

Table 2. *Sample Characteristics of Study 1 (TwinLife)*

	Twin birth cohorts		
	2003/2004	1997/1998	1990-1993
	Numbers		
Twin pairs	1043	1060	984
Male MZ twin pairs	191	218	213
Female MZ twin pairs	230	280	311
Male DZ twin pairs	309	234	199
Female DZ twin pairs	311	327	260
	Age		
Average	11.00	17.01	23.04
Range	10-12	16-18	21-25

Note. MZ: monozygotic; DZ: dizygotic; zygosity is missing for four pairs.

Table 3. *Model Fit Indices for Model Comparisons Regarding Measurement Invariance (MI) across Birth Cohorts of Study 1*

					90% CI			
Trait	Model	χ^2	*df*	RMSEA	LO90	HI90	CFI	ΔCFI
Neuroticism	Unconstrained	204.73	126	.014	.011	.018	.971	
	Metric MI	**212.12**	**130**	**.014**	**.011**	**.018**	**.970**	**.001**
	Strict MI	337.08	142	.021	.018	.024	.929	.042
Extraversion	Unconstrained	269.74	126	.019	.016	.022	.968	
	Metric MI	**293.08**	**130**	**.020**	**.017**	**.023**	**.964**	**.004**
	Strict MI	690.94	142	.035	.033	.038	.879	.089
Openness	Unconstrained	636.67	219	.025	.023	.027	.928	
	Metric MI	**661.00**	**225**	**.025**	**.023**	**.027**	**.921**	**.007**
	Strict MI	950.53	241	.031	.029	.033	.828	.100
Agreeableness	Unconstrained	219.89	126	.016	.012	.019	.975	
	Metric MI	**243.37**	**130**	**.017**	**.014**	**.020**	**.971**	**.004**
	Strict MI	365.63	142	.023	.020	.025	.879	.096
Conscientiousness	Unconstrained	235.13	126	.017	.013	.020	.961	
	Metric MI	**237.67**	**130**	**.016**	**.013**	**.020**	**.961**	**.000**
	Strict MI	491.16	142	.028	.026	.031	.875	.086

Note. *N* = 3,083 twin pairs; the unconstrained model is shown in Figure 2; metric MI: equal factor loadings across cohorts; strict MI: equal factor loadings, error variances, and item residual correlations within same items between twin siblings; best fitting models based on overlapping 90% CIs of RMSEA and ΔCFI < .01 are shown in bold.

Table 4. *Unstandardized Genetic and Environmental Parameter Estimates of the Biometric Structural Equation Model for Study 1 Depicted in Figure 2 and Specific Model-Based Hypothesis Tests*

	Personality traits				
Statistics	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
Latent biometric model parameters					
Genetic effect *g* for cohort 11	0.364	0.315	0.344	0.411	0.393
Genetic effect *g* for cohort 17	0.380	0.560	0.452	0.482	0.575
Genetic effect *g* for cohort 23	0.501	0.569	0.371	0.434	0.513
Environmental effect *e* for cohort 11	0.473	0.468	0.419	0.446	0.598
Environmental effect *e* for cohort 17	0.511	0.572	0.387	0.503	0.551
Environmental effect *e* for cohort 23	0.442	0.539	0.464	0.515	0.436
DZ twins' genetic correlation σ for cohort 11	0.148	-0.011	0.072	0.188	0.620
DZ twins' genetic correlation σ for cohort 17	0.048	-0.021	0.068	0.207	-0.084
DZ twins' genetic correlation σ for cohort 23	0.224	0.060	0.328	-0.001	0.072
Model tests: $\Delta\chi^2$ (Δdf)					
Constrain genetic effects step 1	0.064 (1)	**23.665 (1)**	**5.472 (1)**	1.392 (1)	**9.328 (1)**
Δ*p*	.801	**<.001**	**.019**	.238	**.002**
Constrain genetic effects step 2	**6.376 (1)**	.053 (1)	3.783 (1)	.743 (1)	1.914 (1)
Δ*p*	**.012**	.818	.052	.398	.167
Constrain DZ twins' genetic correlations	.590 (2)	.296 (2)	1.703 (2)	.878 (2)	**7.954 (2)**[a]
Δ*p*	.744	.862	.447	.645	**.019**
Constrain environmental effects	2.634 (2)	**6.689 (2)**[a]	4.752 (2)	1.453 (2)	**12.983 (2)**[b]
Δ*p*	.268	**.035**	.093	.484	**.002**

Note. *N* = 3,083 twin pairs; model parameter estimates are based on the assumption of metric MI; constrain genetic effects step 1: Equal genetic effects across the two younger cohorts $g_{x=11} = g_{x=17}$; constrain genetic effects step 2: Equal genetic effects across the two older cohorts $g_{x=17} = g_{x=23}$; constrain DZ twins' genetic correlations: $\sigma_{x=11} = \sigma_{x=17} = \sigma_{x=23}$; constrain environmental effects: $e_{x=11} = e_{x=17} = e_{x=23}$; significant model tests are shown in bold.

[a]Significant invariance between the youngest cohort and the two older cohorts

[b]Significant invariance between the two younger cohorts and the oldest cohort

Table 5. *Sample Characteristics of Study 2: Norway Twin Study of Youth (NTSY)*

	Study waves		
	2006-2008	2008/2009	2011
		Numbers	
Twin pairs	1,393	1,065	883
Male MZ twin pairs	216	150	122
Female MZ twin pairs	309	245	211
Male DZ twin pairs	208	147	117
Female DZ twin pairs	255	197	174
Opposite-sex DZ twin pairs	405	326	259
Mother raters	1,319	816	390
Father raters	975	578	278
		Age	
Average	15.23	16.92	19.59
Range	12-18	14-21	16-23

Note. MZ: monozygotic; DZ: dizygotic.

Table 6. *Model Fit Indices for Model Comparisons Regarding Measurement Invariance (MI) across Measurement Occasions in Study 2*

					90% CI			
Trait	Model	χ^2	*df*	RMSEA	LO90	HI90	CFI	ΔCFI
Neuroticism	Unconstrained	414.31	300	.016	.012	.019	.982	
	Metric MI	417.75	304	.016	.012	.019	.982	.000
	Strict MI	**433.04**	**312**	**.016**	**.012**	**.019**	**.981**	**.001**
Extraversion	Unconstrained	441.89	300	.018	.014	.021	.983	
	Metric MI	447.14	304	.018	.014	.021	.983	.000
	Strict MI	**476.56**	**312**	**.019**	**.015**	**.022**	**.980**	**.003**
Openness	Unconstrained	489.09	300	.020	.017	.023	.975	
	Metric MI	491.94	304	.020	.017	.023	.975	.000
	Strict MI	**515.00**	**312**	**.021**	**.017**	**.024**	**.973**	**.002**
Agreeableness	Unconstrained	413.01	300	.016	.012	.019	.983	
	Metric MI	421.77	304	.016	.012	.019	.982	.001
	Strict MI	**428.67**	**312**	**.016**	**.012**	**.019**	**.982**	**.001**
Conscientiousness	Unconstrained	404.50	300	.015	.011	.019	.987	
	Metric MI	409.83	304	.015	.011	.019	.987	.000
	Strict MI	**428.04**	**312**	**.016**	**.012**	**.019**	**.986**	**.001**

Note. *N* = 1,393 twin pairs; the unconstrained model is shown in Figure 5; metric MI: equal factor loadings across measurement occasions; strict MI: equal factor loadings, residual variances, and residual correlations across measurement occasions; best fitting models based on overlapping 90% CIs of RMSEA and ΔCFI < .01 are shown in bold.

Table 7. *Unstandardized Genetic and Environmental Parameter Estimates of the Biometric Structural Equation Model for Study 2 Depicted in Figure 5 and Specific Model-Based Hypothesis Tests*

	Personality traits				
Statistics	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
Latent biometric model parameters					
Initial genetic effect g_1 at time 1	0.426	0.437	0.439	0.352	0.413
Initial genetic effect g_{21} at time 2	0.389	0.419	0.397	0.299	0.410
Initial genetic effect g_{31} at time 3	0.407	0.407	0.351	0.306	0.383
Novel genetic effect g_2 at time 2	0.204	0.150	0.210	0.184	0.237
Novel genetic effect g_3 at time 3	0.173	0.172	0.126	0.142	0.198
Novel but stable genetic effect g_{32}	0.146	0.063	0.250	0.146	0.149
Initial environmental effect e_1 at time 1	0.317	0.241	0.171	0.289	0.281
Initial environmental effect e_{21} at time 2	0.304	0.217	0.198	0.255	0.180
Initial environmental effect e_{31} at time 3	0.227	0.164	0.201	0.186	0.096
Novel environmental effect e_2 at time 2	0.157	0.164	0.018	0.106	0.146
Novel environmental effect e_3 at time 3	0.000	0.000	0.000	0.000	0.000
Novel but stable environmental effect e_{32}	0.293	0.282	0.023	0.174	0.216
Model tests: $\Delta\chi^2$ (Δdf)					
Constrain novel genetic effects to zero	**18.852 (3)**	**21.960 (3)**	**49.815 (3)**	**28.640 (3)**	**38.825 (3)**
Δp	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**
Constrain novel environmental effects to zero	**18.272 (3)**	**33.150 (3)**	0.002 (3)	7.445 (3)	**32.560 (3)**
Δp	**<.001**	**<.001**	>.999	.059	**<.001**
Constrain initial/activated genetic effects to be constant	4.508 (3)	**8.281 (3)**	**16.978 (3)**	**9.207 (3)**	**12.349 (3)**
Δp	.212	**.041**	**.001**	**.027**	**.006**
Constrain initial/activated environ. effects to be constant	**8.736 (3)**	**12.222 (3)**	0.771 (3)	**11.364 (3)**	**39.829 (3)**
Δp	**.033**	**.007**	.856	**.010**	**<.001**

Note. N = 1,393 twin pairs; model parameter estimates are based on the assumption of strict MI; constrain novel genetic effects to zero: $g_2 = g_3 = g_{32} = 0$; constrain novel environmental effects to zero: $e_2 = e_3 = e_{32} = 0$; constrain initial/activated genetic effects to be constant: $g_1 = g_{21} = g_{31}$ and $g_2 = g_{32}$; constrain initial/activated environ. effects to be constant: $e_1 = e_{21} = e_{31}$ and $e_2 = e_{32}$; significant model tests are shown in bold.

Table 8. *Heritability and Environmental Components, Genetic and Environmental Stability, and Genetic and Environmental Contributions to the Rank-Order Stability*

	Personality traits				
Statistics	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
Heritability					
Time 1	0.644	0.767	0.868	0.597	0.684
Time 2	0.622	0.728	0.836	0.618	0.807
Time 3	0.612	0.652	0.831	0.676	0.788
Environmental component					
Time 1	0.356	0.233	0.132	0.403	0.316
Time 2	0.378	0.272	0.164	0.382	0.193
Time 3	0.388	0.348	0.169	0.324	0.212
Stability of genetic differences					
Between time 1 and 2	0.886	0.941	0.884	0.852	0.866
Between time 2 and 3	0.920	0.922	0.951	0.917	0.890
Between time 1 and 3	0.874	0.912	0.782	0.832	0.840
Stability of environmental differences					
Between time 1 and 2	0.889	0.798	0.996	0.923	0.777
Between time 2 and 3	0.907	0.922	1.000	0.937	0.891
Between time 1 and 3	0.612	0.503	0.994	0.730	0.406
Genetic contribution to the rank-order stability					
Between time 1 and 2	0.632	0.778	0.837	0.588	0.770
Between time 2 and 3	0.621	0.687	0.827	0.642	0.798
Between time 1 and 3	0.707	0.818	0.818	0.667	0.854
Environmental contribution to the rank-order stability					
Between time 1 and 2	0.368	0.222	0.163	0.412	0.230
Between time 2 and 3	0.379	0.313	0.173	0.358	0.202
Between time 1 and 3	0.293	0.182	0.182	0.333	0.146

Note. Estimates are derived from the biometric structural equation model for Study 2 depicted in Figure 5 and its model parameter estimates shown in the reduced Table 7. See supplementary material for formulas.

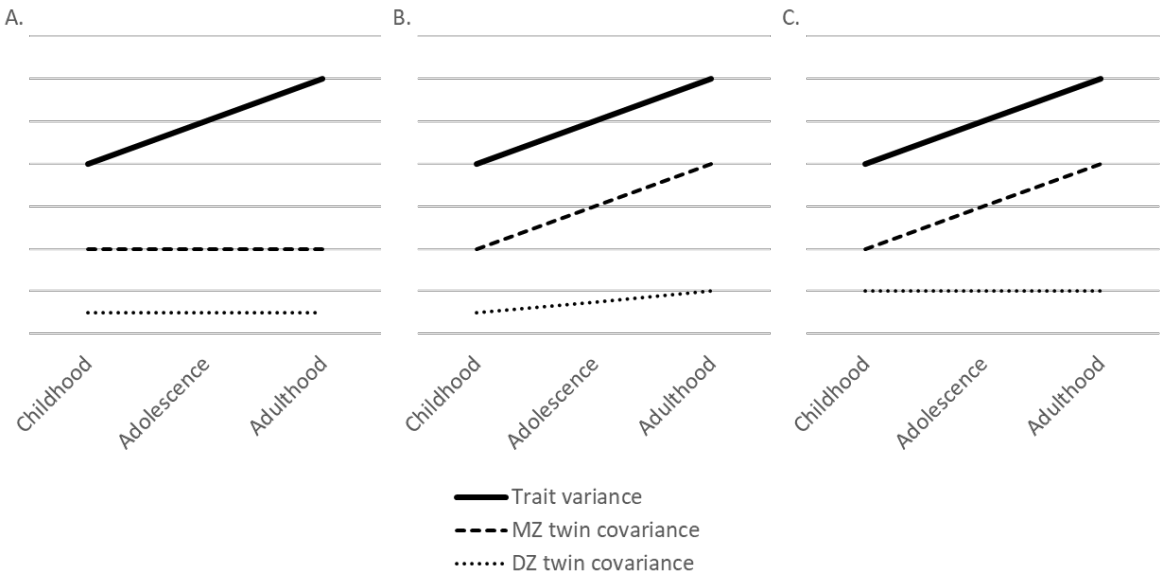

Figure 1. Three example scenarios of potential age trends in the magnitude of trait variance, MZ twin covariance, and DZ twin covariance that have different implications for the underlying sources of increasing trait variance from childhood to adulthood: (A) Environmental individualization account for the increasing trait variance; (B) unfolding of additive and nonadditive genetic factors account for the increasing trait variance; and (C) only nonadditive genetic factors (emergenesis) account for the increasing trait variance. See text and supplementary Table S1 for more details.

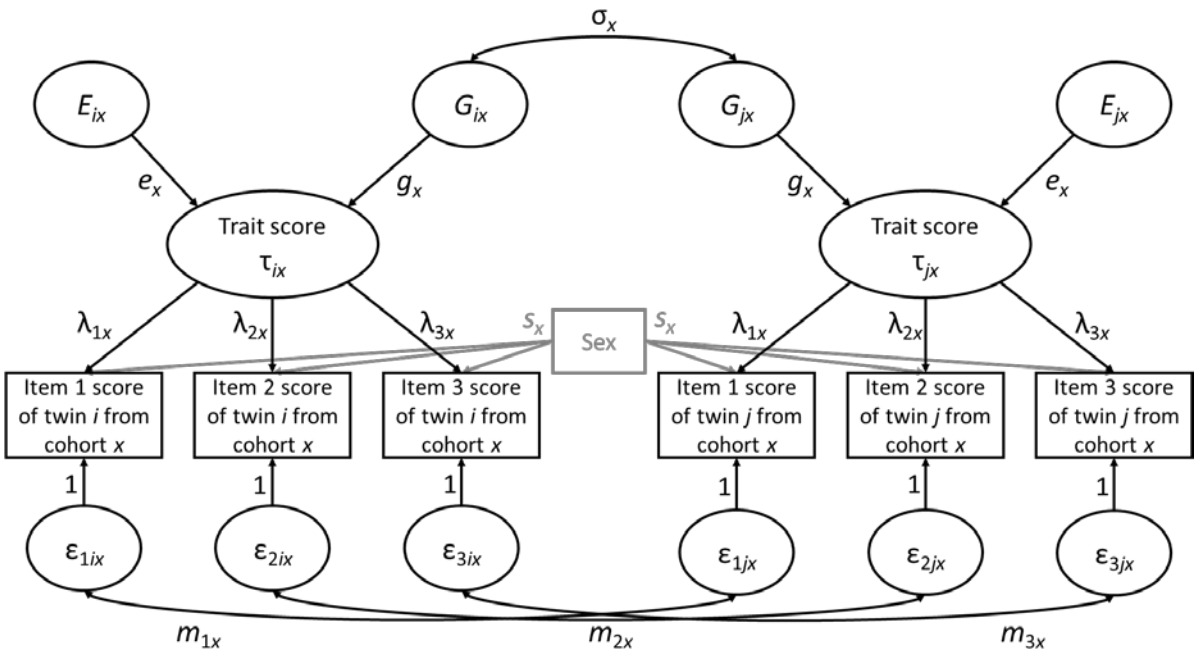

Figure 2. Biometric latent trait score model for Study 1: At the level of measurement model, this SEM allows the decomposition of BFI-item-score variance into latent trait score variance $\lambda^2 \times VAR(\tau)$, variance due to sex differences $s_x \times VAR(\text{Sex})$, and error variance $VAR(\varepsilon)$ for twins *i* and co-twins *j* from birth cohort *x*. Item residuals are allowed to correlate (*m*) within the same item across twin *i* and *j*. At the level of the biometric structure model, the latent trait score variance $VAR(\tau)$ can further be decomposed into a genetic $g^2 \times VAR(G)$ and environmental component $e^2 \times VAR(E)$, provided that variances of latent environmental factors $VAR(E)$ and genetic factors $VAR(G)$ are fixed to 1 in order to estimate e^2 and g^2 as variance components. All parameters are constrained to be equal across twin siblings *i* and *j* as well as across monozygotic (MZ) and dizygotic (DZ) twin pair groups, except the genetic covariance/correlation σ, which is 1 for MZ twins and can take values between 0 and 0.5 for DZ twins. The latter enabled a broad-sense heritability estimate $h^2 = g^2/VAR(\tau)$ including additive and nonadditive genetic differences. All parameters are allowed to vary across cohorts *x*.

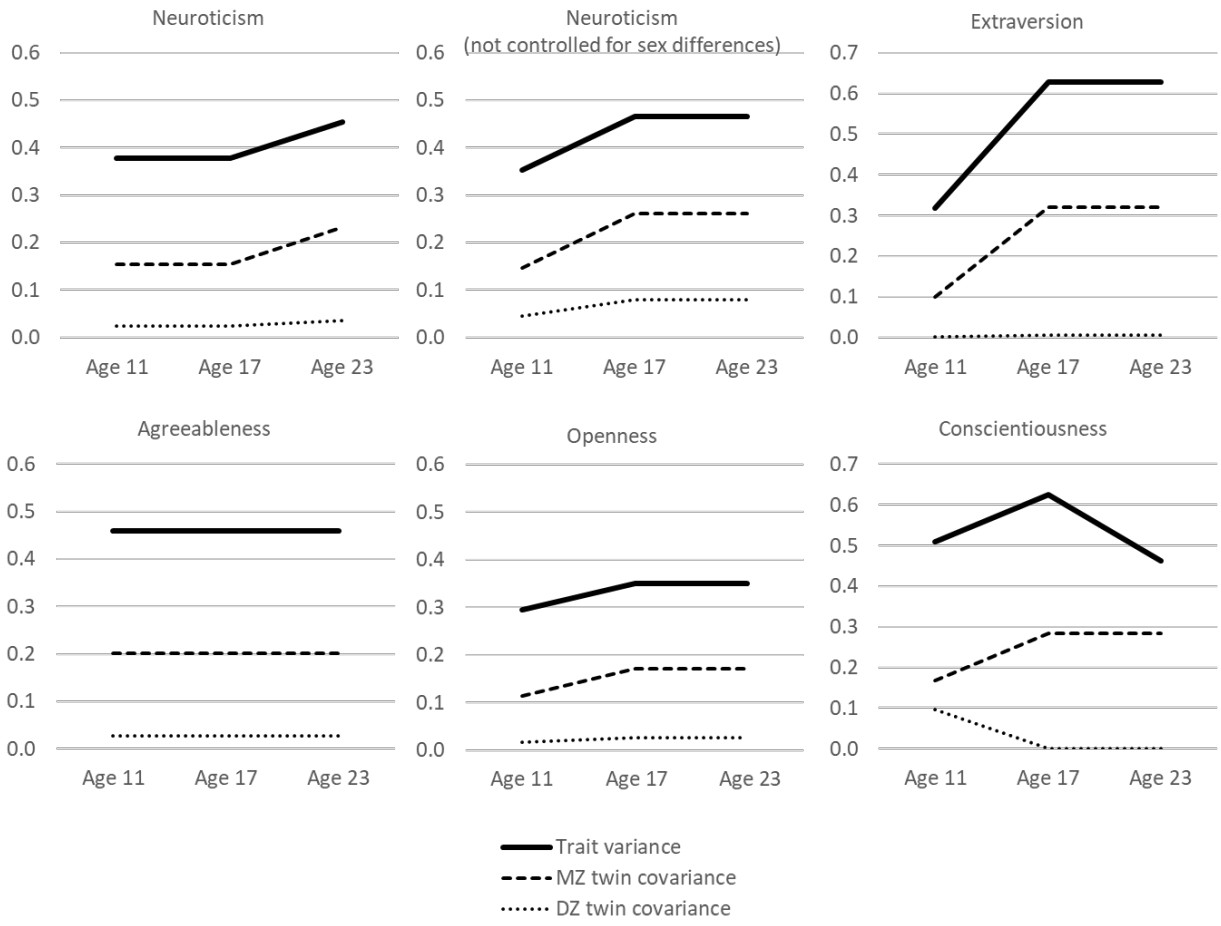

Figure 3. Model-implied age trends of trait variance, MZ twin covariance, and DZ twin covariance for Big Five traits in Study 1. Estimates are based on the most parsimonious biometric models, in which all non-significant differences between cohorts were constrained (see Table S4).

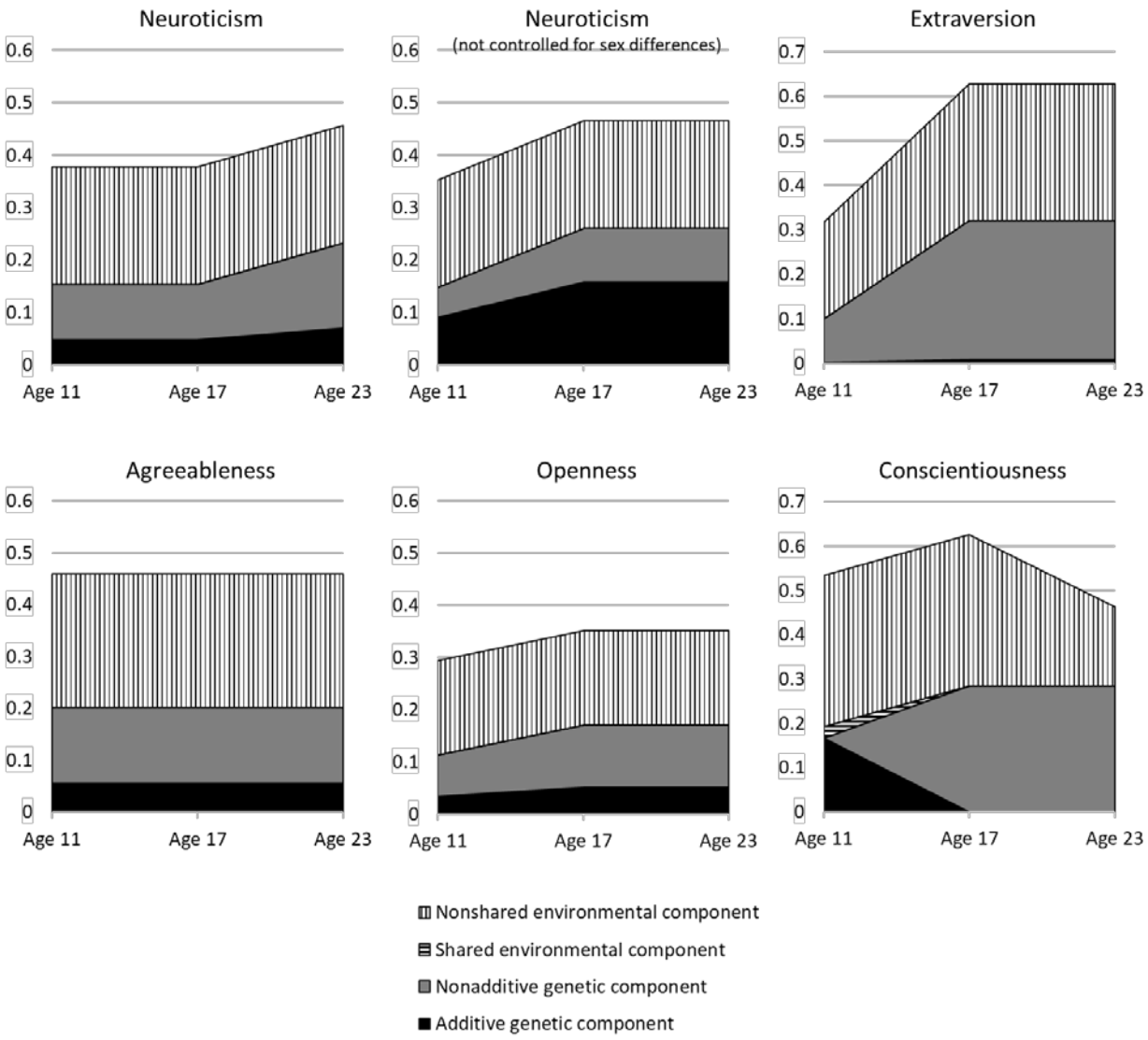

Figure 4. Model-implied age trends of genetic and environmental variance components for Big Five traits in Study 1. Estimates are based on the most parsimonious biometric models, in which all non-significant differences between cohorts were constrained (see Table S4).

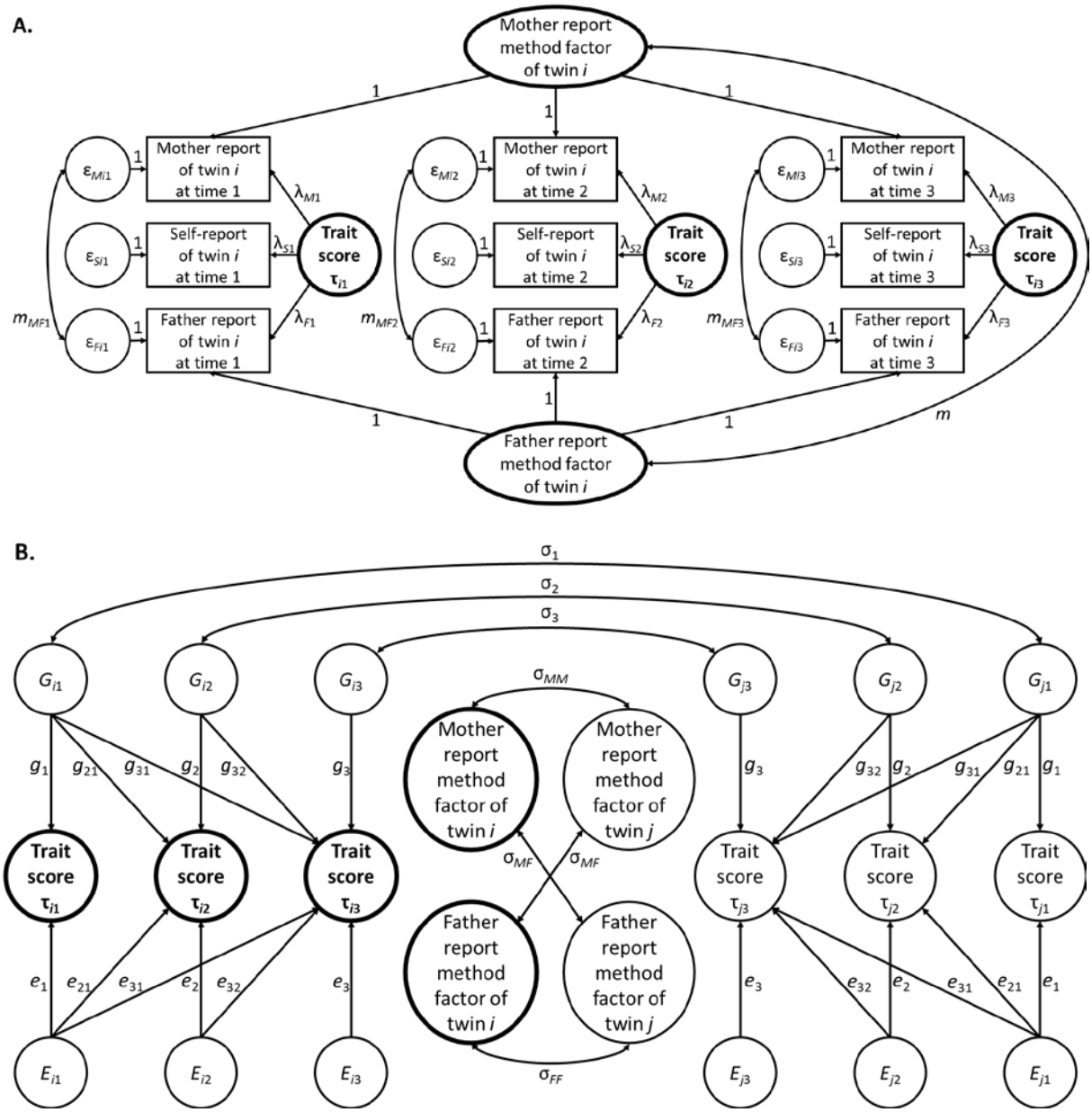

Figure 5. Longitudinal biometric latent trait score model for Study 2: At the level of measurement model (A), this structural equation model allows the decomposition of variance in self-reports and parent ratings into latent trait score variance $\lambda^2 \times VAR(\tau)$ and error variance $VAR(\varepsilon)$ for twins *i* and co-twins *j* (fixing $\lambda_{S1} = \lambda_{S2} = \lambda_{S3} = 1$ for model identification). For simplicity, the measurement model is only shown for twin *i*. Parent report residuals are allowed to correlate within measurement occasions (m_{MF1}, m_{MF2}, and m_{MF3}) and across measurement occasions via correlated method factors (m). Factor loadings and residual variances are allowed to vary across measurement occasions. At the level of the biometric structure model (B), the latent trait score variance $VAR(\tau)$ can further be decomposed into a genetic $g^2 \times VAR(G)$ and environmental component $g^2 \times VAR(E)$, provided that variances of latent environmental factors $VAR(E)$ and genetic factors $VAR(G)$ are fixed to

1 in order to estimate e^2 and g^2. Trait variance at a later measurement occasion (e.g., variance in τ_{i2}) can be accounted for by genetic and/or environmental factors that act to increase trait variance at previous measurement occasions (e.g., g_{21}^2 and e_{21}^2) and novel genetic and environmental factors (e.g., g_2^2 and e_2^2). The different genetic and environmental parameters are constrained to be equal across twin siblings *i* and *j* as well as across monozygotic (MZ) and dizygotic (DZ) twin pair groups, except the genetic covariance/correlation σ, which is 1 for MZ twins and can take values between 0 and 0.5 for DZ twins. The latter enables broad-sense heritability estimate $h^2 = g^2/VAR(\tau)$ including additive and nonadditive genetic differences. Parent-report method factors are also allowed to covary across MZ and DZ twins as a function of twins' genetic relatedness. The factors that appear as model elements in both levels are shown in bold.

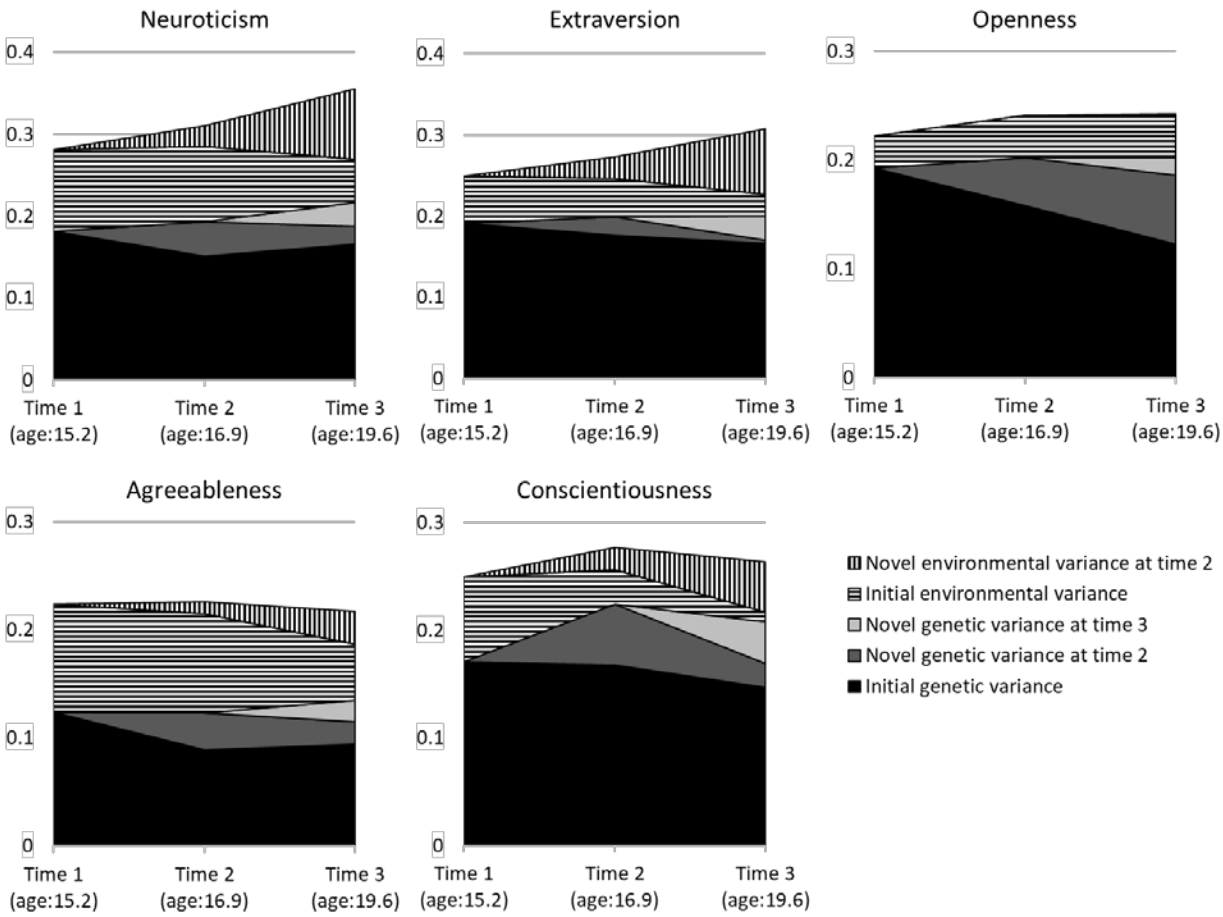

Figure 6. Model-implied age trends of total trait variance and genetic as well as environmental variance components for Big Five personality characteristics in Study 2. Estimates are based on the longitudinal biometric models assuming strict measurement invariance.

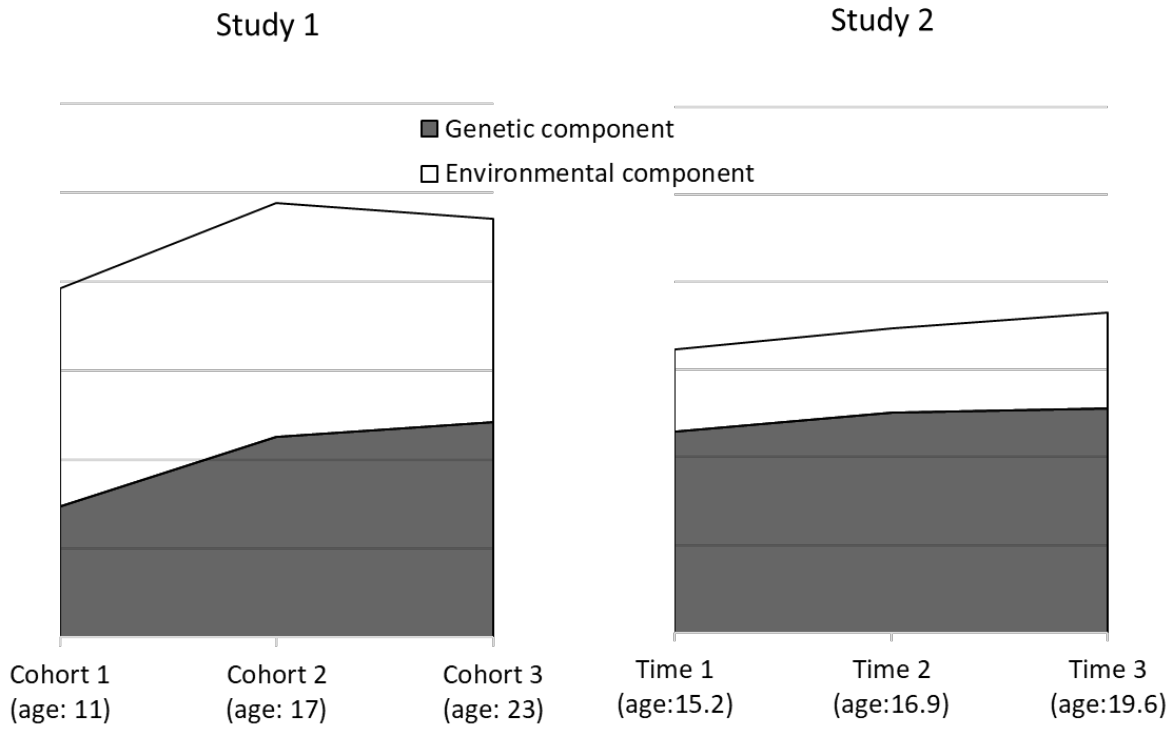

Figure 7. Age trends of total trait score variance and genetic as well as environmental variance components across all Big Five traits based on self-reports in Study 1 and self-ratings as well as parent reports in Study 2.