



Contents lists available at ScienceDirect

Journal of Research in Personality

journal homepage: www.elsevier.com/locate/jrp

Do you see my growth? Two longitudinal studies on personality development from childhood to young adulthood from multiple perspectives [☆]

Ziyan Luan ^{a,*}, Roos Hutteman ^a, Jaap J.A. Denissen ^b, Jens B. Asendorpf ^c, Marcel A.G. van Aken ^a

^a Department of Developmental Psychology, Utrecht University, 3508 TC Utrecht, The Netherlands

^b Department of Developmental Psychology, Tilburg University, 5000 LE Tilburg, The Netherlands

^c Institute of Psychology, Humboldt University Berlin, D-10099 Berlin, Germany

ARTICLE INFO

Article history:

Received 29 November 2015

Revised 24 February 2016

Accepted 23 March 2016

Available online xxx

Keywords:

Personality development

Personality maturation

Disruption hypothesis

Personality judgment

Longitudinal study

Multi-informants

ABSTRACT

Personality developmental studies typically rely on single reporter data, while multi-informant studies are rare. In two longitudinal studies, the present investigation examined inter-judge differences in the development of the Big Five personality traits from childhood to young adulthood. Study 1 investigated personality development as judged by the self and parents from age 12 to 17 to 29 ($N = 186$). Study 2 investigated personality development annually from age 12 to 18 as judged by the self, and both parents and siblings ($N = 574$). Results showed personality maturation from childhood to young adulthood with disruptions during adolescence. Only parent-reports indicated maturation in adolescents' negative affectivity (decreases in N), while self-reports indicated maturation in self-regulatory traits (increases in A and C).

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Personality traits refer to the relatively enduring inter-individual differences in the tendency to feel, think, and behave (Roberts, Wood, & Caspi, 2008).¹ On the one hand, a certain degree of stability is what makes personality traits conceptually distinct from states (Denissen, van Aken, & Roberts, 2011). On the other hand, despite this relative stability, previous research has shown that personality is susceptible to change across the entire life span, especially during young age (e.g., Roberts & DelVecchio, 2000; Roberts, Walton, & Viechtbauer, 2006).

Studies on the development of personality traits have bloomed in the last years (Lucas & Donnellan, 2011; Specht, Egloff, &

Schmukle, 2011; for an overview, see Denissen, 2014). However, the majority of these studies have focused on adulthood, whereas personality development from childhood to young adulthood remains relatively understudied. This is surprising, given that childhood personality predicts a variety of crucial future outcomes, such as parenting (Van den Akker, Deković, Asscher, & Prinzie, 2014), internalizing and externalizing problem behaviors (Denissen, Asendorpf, & van Aken, 2008), and educational and occupational success (Asendorpf, Denissen, & van Aken, 2008). In addition, personality development during childhood and adolescence contains key differences from personality development during adulthood, thus requiring unique scientific attention (Soto & Tackett, 2015).

Most previous studies have relied exclusively on either parent-, teacher-, or self-reports, leaving it unclear whether similar developmental patterns are found when examining personality from multiple perspectives. Cross-sectional studies have shown that judges differ considerably in the information they rely on for personality judgments (Connelly & Ones, 2010; Vazire, 2010). Importantly, there is no single perspective from which a person is known best, rather, both the self and others possess unique information (Vazire & Mehl, 2008). Therefore, multiple informants are needed to capture different perspectives of the developing individuals. The current paper used two longitudinal studies to examine in what way the mean-level change and rank-order stability of the

[☆] Support for this research was provided by the International Society for the Study of Behavioral Development and Jacobs Foundation (ISSBD-JF) Mentored Fellowship Program for Early Career Scholars.

* Corresponding author at: Department of Developmental Psychology, Utrecht University, P.O. Box 80.140, 3508 TC Utrecht, The Netherlands.

E-mail addresses: Z.luan@uu.nl (Z. Luan), roosshutteman@gmail.com (R. Hutteman), jjadenissen@gmail.com (J.J.A. Denissen), asendorpf@gmail.com (J.B. Asendorpf), M.A.G.vanAken@uu.nl (M.A.G. van Aken).

¹ Although this definition could technically also include mental abilities, ability, and temperamental traits - which were historically covered in relatively separate literature - in the current paper we focus on the development of personality traits.

Big Five personality traits from childhood to early adulthood differ depending on the judge,² and the level of self-other agreement.

Previous studies have shown many more substantial changes in personality from childhood to young adulthood compared to the later ages, reflected in both rank-order stability and mean-level change. Rank-order stability reflects whether groups of people maintain their relative placement to each other on personality traits over time. A classic meta-analysis based on 152 longitudinal studies showed that rank-order stabilities were moderate during early childhood and adolescence, and large from college years to old age (Roberts & DelVecchio, 2000). A more recent meta-analysis (Ferguson, 2010) confirmed the significantly lower rank-order stability from childhood to young adulthood, and further recommended the consideration of measurement error when investigating rank-order stability.

Another type of change - mean-level change - reflects the average amount of change in the population as a whole, independent of individual differences. A meta-analysis of 92 longitudinal studies (Roberts et al., 2006) showed that people, on average, increased in social dominance (a facet of extraversion) and conscientiousness and decreased in neuroticism, especially during young adulthood (age 20–40). Moreover, people increased on social vitality (another facet of extraversion) and openness in adolescence, but then decreased in both of these domains during old age. Agreeableness showed no mean-level change until old age (after age 50), when it increased.

Recent theoretical frameworks have aimed to describe developmental patterns in these results. The “maturity principle” refers to the finding that individuals tend to become more conscientious, more agreeable, and less neurotic with age (Bleidorn et al., 2013; Roberts et al., 2008). However, the maturity principle was based on findings focusing on adults, and more recent studies have shown that personality development during adolescence is more in accordance with the disruption hypothesis (Denissen, van Aken, Penke, & Wood, 2013; Soto & Tackett, 2015). The disruption hypothesis suggests that the biological, social, and psychological transitions from childhood to adolescence are accompanied by temporary dips in some aspects of personality maturity, thus showing a temporary deviation from the maturity principle during adolescence (Denissen et al., 2013; Klimstra, Hale, Raaijmakers, Branje, & Meeus, 2009; Van den Akker et al., 2014).

Although these recent studies have provided valuable insights into developmental patterns of personality during young age, they have typically relied on single-reporter data, while multi-informant studies are rare. However, cross-sectional studies have shown that judges differ considerably in their judgment of personality traits. The Self-Other Knowledge Asymmetry Model (the SOKA Model; Vazire, 2010) advocates that judges vary considerably in their information and motivation for personality judgments. Therefore, personality judgment might be, at least to some degree, a social construction. Transferring this to a developmental framework, children’s personality maturation and the possible disruption of this maturation during adolescence might be observed differently by different judges.

The constructivist perspective and the realistic perspective, regardless of their different assumptions of the degree to which “the true” personality exists, both provide support for this notion. Studies from a more constructivist perspective maintain that alternative personality judgments are both valid, since each reflects accurately what this judge perceives (e.g., John & Robins, 1993). Studies from a more realistic perspective maintain that valid cues

need to be available and used, in order to make accurate personality judgment. However, the availability and usage of valid cues are almost always not perfectly sufficient, and then personality judgments are influenced by various heuristics (e.g., Funder, 1995), such as convenient social comparisons (Wood, Brown, Maltby, & Watkinson, 2012) or current relationship quality (Watson, Hubbard, & Wiese, 2000).

Studying personality development from multiple perspectives is important, because recent studies have shown that there is no single perspective from which a person is known best. Rather, both the self and others possess unique information (e.g., Vazire & Mehl, 2008). In addition, the perceived views of each other’s personalities influence the interpersonal interaction and as such thus deserve more scientific attention. However, the notion that judges might differ in the degree to which or the personality trait in which they observe personality maturation and possible disruptions thereof, has rarely been tested longitudinally.

A highly interesting exception by Watson and Humrichouse (2006) tracked newlywed young adults for two years, and found that while self-ratings were in accordance with the maturity principle – increases in conscientiousness and agreeableness and decreases in neuroticism over time – spouses reported opposite developmental trajectories of the very same person’s personality, specifically decreases in conscientiousness, agreeableness, extraversion, and openness.

How can these results be translated to differences between judges when looking at personality development in childhood and adolescence? When focusing on the most important relationship partners during childhood – the parents – previous studies have shown that parents possess some of the characteristics of “good judges” in that they are motivated to provide thoughtful responses about their child and are highly familiar with their child (Funder, 1995; Tackett, Herzhoff, Kushner, & Rule, 2015). However, just like other judges, a parent’s judgment of their child’s personality and emotions can also be biased (Durbin & Wilson, 2012; Tackett, 2011). Consequently, mothers’ and fathers’ longitudinal judgments of their children’s personality might differ from each other, and also differ from the judgments of children themselves and of other family members.

Indeed, a recent study by Van den Akker et al. (2014) investigated personality development by self- and mother-report and found that benevolence and conscientiousness increased from middle to late childhood, temporarily declined from late childhood to mid-adolescence, and increased again thereafter. Imagination decreased from middle childhood to mid-adolescence and also increased again thereafter. Mothers reported a temporary decline in emotional stability, which was not confirmed by children’s self-ratings.

A number of questions still remain unknown in this field of research. First, within the family context, personality judgments by fathers and siblings are also important to understand the development of an adolescent’s personality. Fathers and mothers show only moderately high agreement regarding their child’s personality traits (Tackett, 2011). Moreover, sibling relationships are among the most constant and prominent social companionships in adolescence (Jenkins & Dunn, 2009). Adolescent siblings are of similar age and encounter the same developmental tasks and emotional fluctuations, therefore both mean-level change and rank-order stabilities of siblings’ personality judgments might be more similar to adolescents’ self-views than to parents’ views.

Second, the study by Van den Akker et al. (2014) compared personality development judged by children and mothers from age 9 to 17. It would be interesting to see whether parent-ratings confirm the maturity principle in the longer term, after the “storm and stress” period of adolescence (Arnett, 2000; Casey et al., 2010).

² There are multiple ways of referring to who judges the personality, such as the judge, rater, perceiver, and reporter. In the current paper we will consistently use the term “judge”.

Third, not much is known regarding the self-other agreement in personality during childhood and adolescence. On the one hand, lay people show wisdom in personality judgments. Previous studies have shown that people in general know what cues are valid for personality judgments and actually use these valid cues to form their personality judgments (Funder & Sneed, 1993). One of the pioneer studies by Funder and Colvin (1988) found, on average, medium level agreement between the self and close friends across personality traits. Moreover, people know that others see their personality differently from how they see themselves, and they have a pretty good idea about the impressions they make to others (Carlson & Furr, 2009; Carlson, Vazire, & Furr, 2011). On the other hand, inter-judge differences in personality are also considerable, and each perspective provides unique predictive validities (Connelly & Ones, 2010; Vazire & Mehl, 2008).

The SOKA Model (Vazire, 2010) also suggests that more visible and behaviorally centered traits (i.e., extraversion and conscientiousness) should be more consistently judged than less visible and evaluative traits (i.e., neuroticism, openness, and agreeableness). Cross-sectional studies confirm this expectation for adults, but it remains unclear whether previous findings from adult research on the SOKA Model generalize to personality traits in childhood and adolescence. In addition, it is questionable whether the *development* of visible traits, such as extraversion and conscientiousness, is also more consistently judged than the *development* of the other traits.

2. The present study

The present investigation aimed to examine differences in the development of the Big Five personality traits in childhood, adolescence, and young adulthood across judges (i.e., self-, mother-, father-, and sibling-ratings) in two longitudinal studies. Study 1 examined the personality development of German children from childhood to young adulthood (age 12 to 17 to 29; $N = 155$) as judged by the children themselves as well as their parents. Study 2 zoomed in on adolescence, to examine personality development from age 11.5 to 17.5, assessed annually, with an accelerated longitudinal design ($N = 576$ Dutch adolescents).

Previous work by Branje, van Lieshout, and Gerris (2007) based on the same dataset with half of the current sample, found first indications that mean-level change in personality during adolescence might differ between self-ratings and aggregated other-ratings (i.e., aggregating the personality judgments from three family members). The current paper moved beyond this by examining specific differences between judges (i.e., self, father, mother, and sibling) in judging personality development in adolescence, shown through mean-level change, rank-order stability, and self-other agreement. We had three research questions.

2.1. Is personality maturation from childhood to young adulthood perceived differently across judges?

We examined the maturity principle (Roberts et al., 2006) from childhood to young adulthood. Based on the seemingly robust and universal findings for the maturity principle obtained from meta-analyses and cross-cultural studies (Bleidorn et al., 2013; Roberts et al., 2008), we expected that with a relatively long time interval (Study 1), both self and parent-reports of personality would be in line with the maturity principle (i.e., increases in agreeableness and conscientiousness, and decreases in neuroticism). Due to limited current knowledge, we formulated no specific hypotheses concerning the question of whether children and their parents would see similar degrees of personality maturation.

2.2. Are personality maturation and disruption during adolescence perceived differently across judges?

Zooming in on adolescence (Study 2) we examined the possible dips in some aspects of personality maturity during adolescence, as advocated by the disruption hypothesis (Denissen et al., 2013; Soto & Tackett, 2015). We expected that during adolescence, there would be some developmental trajectories that deviated from the maturity principle (i.e., decreases in agreeableness and conscientiousness, or increases in neuroticism), as observed by at least some of the judges. Given the increases in parent-adolescent conflict during adolescence (Van der Giessen et al., 2014), we expected parents to observe more disruptions in personality maturation than adolescents themselves or their siblings observed. Due to the limited current knowledge, we did not formulate a hypothesis regarding whether all judges would observe disruptions in personality maturation.

2.3. Does the SOKA model apply to children and adolescents in a developmental framework?

We examined whether there would be higher self-other agreement in more visible and behavioral oriented traits than less visible and more evaluative traits during childhood and adolescence. In line with the SOKA Model (Vazire, 2010), we expected that (in both studies) the self-other agreement for extraversion and conscientiousness would be higher than for other personality traits, such as neuroticism. In addition, we examined whether the *development* of more visible traits would also be more consistently judged than the *development* of less visible and more evaluative traits from multiple-perspectives. In line with the SOKA Model, we expected that (in both studies) both the mean-level change and rank-order stability of extraversion and conscientiousness would be more consistently judged than of the other personality traits, such as neuroticism.

3. Study 1

Study 1 investigated longitudinally personality development from childhood to adolescence to young adulthood (age 12 to 17 to 29) as judged by children themselves as well as their parents.

3.1. Method

3.1.1. Participants and procedure

Participants were part of the Munich Longitudinal Study on the Genesis of Individual Competencies (LOGIC; Weinert & Schneider, 1999). The first wave started in the fall of 1984 in the Munich area. The LOGIC sample initially contained 230 children (119 boys) who started preschool in the Munich area at the age of 3 or 4 years old. Their first language was German. Twenty schools were selected from a broad spectrum of neighborhoods, and more than 90% of parents who were asked, gave consent for their child's participation. The present study included three waves of measurements, when participants were on average 12 years old (186 self-ratings and 173 parent-ratings – mainly mother-ratings, tested in 1992), 17 years old (174 self-ratings and 146 mother-ratings, tested in 1998) and 29 years old (153 self-ratings and 81 mother-ratings, tested in 2010). Although participants were also examined at age 23, those data were not included due to a change of the personality measure at that wave.

Attrition analyses revealed that for all of the investigated variables, there were no significant differences between complete and incomplete cases. More specifically, we conducted a one-way ANOVA to compare complete cases (i.e., cases that showed no

missingness for all research variables, $N = 64$) and cases with missingness (i.e., cases that showed missingness for at least one research variable, $N = 122$), on all research variables (at parcel level). Results showed that for all research variables, the 95% CI of complete cases overlapped with the 95% CI of cases with missingness.

In addition, we conducted post-hoc power analyses with Monte Carlo simulation studies in Mplus Version 7.11 (Muthén & Muthén, 2013) following Muthén and Muthén (2002), based on the estimated parameter value and missingness obtained from the present study. The number of replications was set to 5000 to achieve a stable estimation. Regarding the mean-level change of self- and parent-ratings, results showed that for all Big Five personality traits, power was higher than .80 for investigating both intercepts and slopes, indicating sufficient statistical power.

Regarding the rank-order stability of self- and parent-ratings, results showed that for all Big Five personality traits, power was higher than .80 for investigation of rank-order stability during both time intervals (i.e., age 12–17, and age 17–29). However, it should be noted that for models of mean-level change and rank-order stability, there were warnings (e.g., lack of convergence) in some of the 5000 replication cases, indicating that although statistical power was sufficient for our study, it would be more optimal to have a larger sample size. In addition, because inter-judge differences in mean-level change and rank-order stability were tested by chi-square differences tests - the power estimation of which is currently beyond the capability of Mplus - it is possible that with a larger sample size, additional smaller inter-judge differences could be captured.

Regarding self-other agreement, results showed that for all but three cases of the Big Five personality traits at all measurement waves, power was higher than .80 for investigating self-parent agreement during the three waves. The three exceptions were self-parent agreement in neuroticism and openness at age 17 (neuroticism: $r = 0.13$, power = 0.30; openness: $r = 0.21$, power = 0.56), and self-parent agreement in agreeableness at age 29 ($r = 0.16$, power = 0.45), when self-parent agreement was very low.

3.1.2. Measures

Big Five personality traits. Participants and their parents provided ratings on neuroticism, agreeableness, conscientiousness, extraversion, and openness using bipolar adjective pairs that were obtained from Ostendorf (1990). Both children and parents rated the items on a 5-point scale (from 1 = *totally agree with the adjective word on the left side*, to 5 = *totally agree with the adjective word on the right side*). Sample items include for neuroticism: calm vs. irritable; for agreeableness: vengeful vs. forgiving; for conscientiousness: lazy vs. diligent; for extraversion: unsocial vs. outgoing; and for openness: uneducated vs. knowledgeable.

The eight items of every Big Five dimension were parceled into three indicators per dimension. We applied the Item-to-Construct Balancing approach (Little, Cunningham, Shahar, & Widaman, 2002), in which items with the highest loadings were used to anchor the three parcels. Subsequently, the items with the next highest factor loadings were added to the anchor items in inverted order until all items were assigned to a parcel. The same parceling structure was applied for self-ratings and parent-ratings at all waves. More specifically, we first applied the Item-to-Construct Balancing procedure for self-rating at the first measurement wave to achieve a parceling structure. Second, this parceling structure was revised based on CFA results of all the judges at all waves, aiming to maximize the possibility that for all the judges at all waves each parcel would show a similar factor loading. Third, the final parceling structure was applied to all the judges at all waves to ensure the comparability. As shown in Table 1, Cronbach's Alphas were satisfactory.

Table 1
Study 1 and Study 2: Cronbach's Alphas of the Big Five personality traits.

| | Judge | N | A | C | E | O |
|---------|---------|---------|---------|---------|---------|---------|
| Study 1 | Self | .74–.88 | .77–.83 | .82–.91 | .84–.90 | .67–.86 |
| | Parent | .82–.85 | .84–.87 | .90–.91 | .89–.89 | .85–.91 |
| Study 2 | Self | .64–.86 | .72–.85 | .73–.92 | .64–.89 | .58–.85 |
| | Mother | .80–.89 | .81–.92 | .91–.96 | .87–.93 | .82–.89 |
| | Father | .76–.88 | .82–.90 | .90–.94 | .83–.93 | .80–.89 |
| | Sibling | .68–.88 | .76–.92 | .81–.93 | .75–.90 | .62–.81 |

Note. N: Neuroticism, A: Agreeableness, C: Conscientiousness, E: Extraversion, O: Openness.

3.1.3. Analytic strategy

Missing data handling and model fit. Missing data were handled using full information maximum likelihood (FIML) estimation, thereby making optimal use of the available data. Model fit was assessed using the comparative fit indices (CFIs) and root-mean-square error of approximation (RMSEA). CFI values of .90 and higher and RMSEA values of .08 and lower reflect an acceptable fit to the data (Marsh, Hau, & Grayson, 2005).

Mean-level change in self- and parent-ratings of the Big Five personality traits. We first conducted a multiple-group CFA with the specification of measurement invariance across waves and across judges. Children themselves and their parents were specified as two groups. Different levels of measurement invariance (MI) can be achieved: Weak MI requires only identical factor loadings across time/judges, strong MI requires additionally that intercepts be invariant across time/judges, and strict MI requires invariant residual variance in addition to the invariant factor loadings and intercepts. In all analyses, we specified the strictest possible measurement invariance for good model fit.

Mean-level changes were tested by adding the estimation of intercepts and slopes of the latent trait from age 12 to age 29 to the multiple-group CFA models. All three waves had loadings of 1 on the intercept variable. For the slope, age 12 had a loading of 0, age 29 had a loading of 1, and age 17 was freely estimated. Whether the differences in intercepts and slopes between judges were significant, was tested by Chi-square difference tests. A significant decline in model fit when constraining the parameter estimations across groups to be equal would indicate differences between judges, whereas a lack of significant change in model fit would indicate no differences between judges in mean-level change.

Rank-order stability in self- and parent-ratings of the Big Five personality traits. We first conducted a multiple-group CFA with the specification of measurement invariance. Rank-order stability was investigated by adding the estimation of the correlation of the latent trait between age 12 and 17, as well as the estimation between age 17 and 29 to the multiple-group CFA model. Whether differences in the correlation of a latent trait over time between judges were significant or not was tested by Chi-square difference tests. Two time lags (age 12–17, and 17–29) were tested one by one.

Self-parent agreement in the Big Five personality traits. We conducted a CFA with the specification of measurement invariance for each of the Big Five personality traits separately. Self-parent agreement in a certain personality trait refers to the correlation coefficients between the two latent factors (i.e., self-rating and parent-rating).

3.2. Results

Means, standard deviations and inter-correlations of all manifest variables can be found in Table S2 in the online Supporting Information. The results concerning the mean-level change of the

Big Five personality traits as judged by children themselves and their parents are reported first, followed by the results for the rank-order stability of the Big Five personality traits. Finally, self-parent agreements for the Big Five personality traits from age 12 to 29 are reported.

3.2.1. Mean-level development of self- and parent-ratings of the Big Five personality traits

Model fit indices and the estimated mean-level changes of the Big Five personality traits from age 12 to age 29 are shown in Table 2 and Fig. 1.³

Neuroticism. Parents on average judged their children to be more neurotic than the children judged themselves to be at age 12 ($\Delta\chi^2 = 24.80$, $\Delta df = 1$, $p < .001$; Cohen's $d = 0.76$). From age 12 to 29, children on average judged themselves to be stable on neuroticism, whereas parents observed a significant decrease in their children's neuroticism ($M_{parent} = -0.34$, $p < .001$, 95% CI [-0.48, -0.20]; $M_{self} = -0.08$, $p = .171$, 95% CI [-0.21, 0.04]; parent vs. self: $\Delta\chi^2 = 8.13$, $\Delta df = 1$, $p = .004$; Cohen's $d = 0.57$).

Agreeableness. Children and their parents observed similar levels of agreeableness at age 12 ($\Delta\chi^2 = 0.92$, $\Delta df = 1$, $p = .337$, Cohen's $d = 0.16$). From age 12 to 29, children and their parents observed similar increases in children's agreeableness ($M_{self} = 0.14$, $p < .001$, 95% CI [0.06, 0.21]; $M_{parent} = 0.15$, $p = .019$, 95% CI [0.03, 0.27]; parent vs. self: $\Delta\chi^2 = 0.05$, $\Delta df = 1$, $p = .823$).

Conscientiousness. Children judged themselves to be more conscientious than their parents judged them to be at age 12 ($\Delta\chi^2 = 8.73$, $\Delta df = 1$, $p = .003$; Cohen's $d = 0.43$). From age 12 to 29, although both self- and parent-ratings increased significantly, parents reported a stronger increase in conscientiousness than the children did ($M_{self} = 0.51$, $p < .001$, 95% CI [0.42, 0.61]; $M_{parent} = 0.75$, $p < .001$, 95% CI [0.60, 0.90]; parent vs. self: $\Delta\chi^2 = 8.02$, $\Delta df = 1$, $p = .005$).

Extraversion. Parents judged their children's extraversion at age 12 at the same level as children judged themselves ($\Delta\chi^2 = 0.33$, $\Delta df = 1$, $p = .566$; Cohen's $d = 0.08$). From age 12 to 29, extraversion was stable, independent of who judged ($M_{self} = -0.08$, $p = .063$, 95% CI [-0.16, 0.00]; $M_{parent} = 0.13$, $p = .071$, 95% CI [-0.01, 0.27]).

Openness. Parents on average judged their children to be more open to experience at age 12 than children judged themselves ($\Delta\chi^2 = 6.31$, $\Delta df = 1$, $p = .012$; Cohen's $d = 0.35$). From age 12 to 29, although children's increase in openness was reported by both themselves and their parents, parents reported a greater increase ($M_{self} = 0.15$, $p < .001$, 95% CI [0.07, 0.23]; $M_{parent} = 0.45$, $p < .001$, 95% CI [0.35, 0.55]; parent vs. self: $\Delta\chi^2 = 20.52$, $\Delta df = 1$, $p < .001$; Cohen's $d = 0.80$).

3.2.2. Rank order stability of self- and parent-ratings of the Big Five personality traits

Rank-order stabilities of the Big Five personality traits from age 12 to age 29, as well as the model fit indices, are shown in Table 3 and Fig. 2.

Models with identical stabilities between self- and parent-ratings did not significantly worsen model fit for any of the personality traits except for neuroticism between age 12 to age 17, in which self-ratings showed a lower stability than parent-ratings ($r_{self} = 0.25$, $p = .025$; $r_{parent} = 0.67$, $p < .001$; parent vs. self: $\Delta\chi^2 = 12.86$, $\Delta df = 1$, $p < .001$). According to Cohen's criterion (1992), a correlation coefficient of 0.1 indicates a small effect size, 0.3 indicates a medium effect size, and 0.5 indicates a large effect size. In Study 1, self-ratings showed small to medium rank-order stability from age 12 to 17, while parent-ratings showed large

rank-order stability. In sum, results indicated similar developmental patterns of rank-order stability of personality traits with the exception of neuroticism from age 12 to age 17, which was significantly less stable in the eyes of the children than in the eyes of their parents.

3.2.3. Self-parent agreement in the Big Five personality traits

As shown in Fig. 3, on average children and their parents showed highest agreement in conscientiousness, followed by extraversion, while agreement was lowest for neuroticism and openness. Model fits for the Big Five personality traits were all satisfactory: for neuroticism: $\chi^2(145) = 256.16$, $p < .001$; CFI = .91, RMSEA = .06; for agreeableness: $\chi^2(145) = 199.26$, $p = .002$; CFI = .96, RMSEA = .05; for conscientiousness: $\chi^2(145) = 284.14$, $p < .001$; CFI = .93, RMSEA = .07; for extraversion: $\chi^2(145) = 306.64$, $p < .001$; CFI = .91, RMSEA = .08; for openness: $\chi^2(143) = 260.93$, $p < .001$; CFI = .92, RMSEA = .07.

In addition, during adolescence self-parent agreement in neuroticism dropped considerably, from medium to large agreement at age 12, to non-significant and small agreement at age 17. Self-parent agreement in neuroticism rebounded to medium level by age 29. Another noteworthy change of self-parent agreement was agreeableness, which showed medium to high agreement from age 12 to 17, yet at age 29 self-parent agreement dropped to non-significant.

3.3. Discussion

Study 1 longitudinally investigated personality development from age 12 to 29, as judged by children themselves as well as their parents. We discuss the results below in accordance with our research questions.

3.3.1. Is personality maturation from childhood to young adulthood perceived differently across judges?

First, we examined the maturity principle and disruption hypothesis from childhood to young adulthood as judged by children themselves as well as their parents. In line with our expectations, even with this relatively long time interval, the overall developmental trend as perceived by both judges was bright. That is, both self- and parent-ratings of personality development from childhood to young adulthood were supportive of the "maturity principle" (Roberts et al., 2006).

Second, we examined the hypothesis that the maturity principle is a social construction that is observed differently by different judges (Watson & Humrichouse, 2006). Supporting this notion, results showed that parents observed stronger personality maturation in their children than the children did themselves. Specifically, parent-ratings showed a comparable amount of increase in agreeableness, but a greater increase in conscientiousness than the self-ratings. Similarly, parents' ratings indicated a significant decrease in children's neuroticism, whereas the children's ratings of their own neuroticism were stable. These results indicate that parents have a more positive view of their children's personality development from childhood to young adulthood than their children themselves have.

3.3.2. Does the SOKA model apply to children and adolescents in a developmental framework?

First, we examined the validity of the SOKA model in childhood and adolescence. In line with our expectations, children at the beginning of adolescence tended to judge their own personality as more conscientious, less neurotic, but also less open to experience than their parents judged their personality. In addition, results showed that more visible and behavior-oriented traits (i.e., conscientious and extraversion) showed higher self-parent

³ Because of the long time-interval and the possible disruptions in between the measurements, the imaginary lines are only rough indications of developmental trajectories and should be interpreted with caution.

Table 2
Study 1: Mean-level change of self- and parent- ratings of the Big Five personality traits.

| Model | χ^2 (df) | CFI | RMSEA | Judge | Means of intercept | Variance of intercept | Means of slope | Variance of slope |
|-------|----------------|-----|-------|--------|--------------------|-----------------------|----------------|-------------------|
| N | 149.29 (69)*** | .93 | .08 | Self | 2.22*** | 0.00 | -0.09 | 0.22** |
| | | | | Parent | 2.52*** | 0.31*** | -0.34*** | 0.17 |
| A | 102.44 (70)** | .97 | .05 | Self | 3.72*** | 0.08*** | 0.14*** | 0.00 |
| | | | | Parent | 3.77*** | 0.13*** | 0.15*** | 0.00 |
| C | 147.96 (67)*** | .95 | .08 | Self | 3.29*** | 0.19*** | 0.51*** | 0.00 |
| | | | | Parent | 3.07*** | 0.33*** | 0.75*** | 0.00 |
| E | 158.71 (76)*** | .95 | .08 | Self | 3.89*** | 0.20*** | -0.08 | 0.00 |
| | | | | Parent | 3.85*** | 0.31*** | 0.13 | 0.00 |
| O | 130.68 (59)*** | .95 | .08 | Self | 3.92*** | 0.04* | 0.15*** | 0.13*** |
| | | | | Parent | 4.04*** | 0.19*** | 0.45*** | 0.15*** |

Note. N: Neuroticism, A: Agreeableness, C: Conscientiousness, E: Extraversion, O: Openness. CFI: Comparative Fit Index; RMSEA: Root Mean Square Error of Approximation.

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

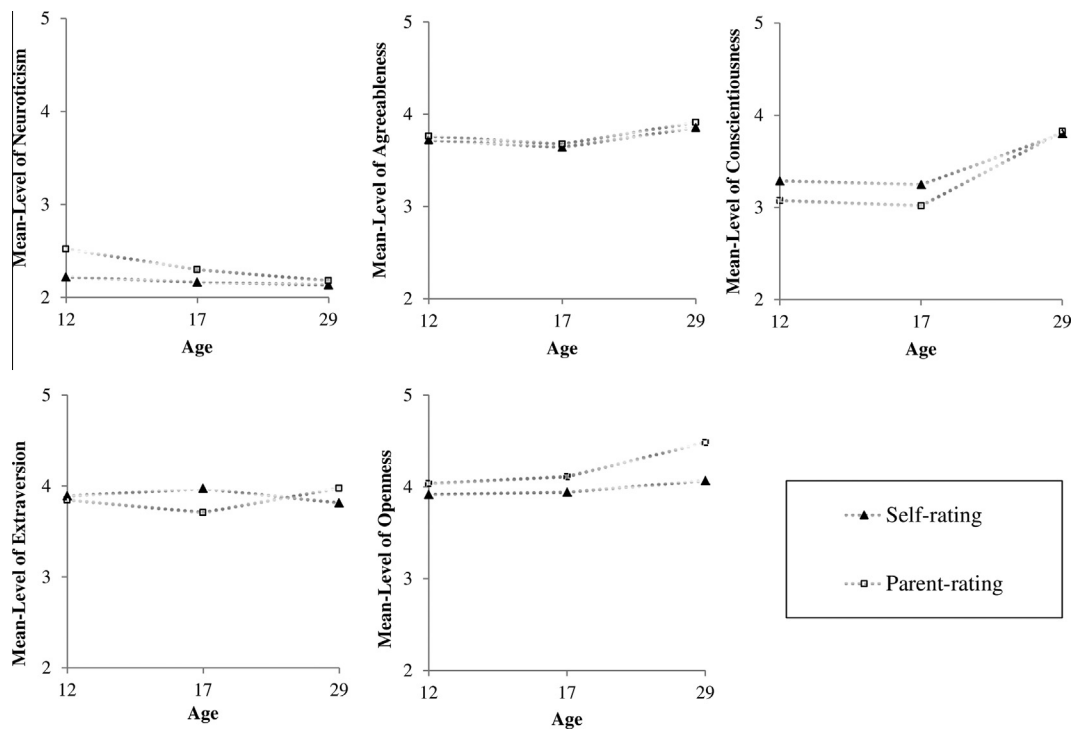


Fig. 1. Mean-level change of the Big Five personality traits from age 12 to 29.

Table 3
Study 1: Rank-order stabilities of self- and parent-ratings of the Big Five personality traits.

| Model | χ^2 (df) | CFI | RMSEA | Judge | Stability T1 (age 12–17) | S.E. | Stability T2 (age 17–29) | S.E. |
|-------|---------------|-----|-------|--------|--------------------------|------|--------------------------|------|
| N | 168.08(76)*** | .92 | .08 | Self | .25* | 0.11 | .47*** | 0.08 |
| | | | | Parent | .67*** | 0.06 | .47*** | 0.10 |
| A | 99.35(70)* | .97 | .05 | Self | .56*** | 0.07 | .42*** | 0.09 |
| | | | | Parent | .53*** | 0.07 | .23 | 0.12 |
| C | 173.47(76)*** | .95 | .08 | Self | .57*** | 0.06 | .65*** | 0.06 |
| | | | | Parent | .57*** | 0.06 | .68*** | 0.07 |
| E | 174.60(80)*** | .95 | .08 | Self | .67*** | 0.05 | .49*** | 0.07 |
| | | | | Parent | .57*** | 0.06 | .54*** | 0.08 |
| O | 146.42(65)*** | .94 | .08 | Self | .44*** | 0.10 | .54*** | 0.07 |
| | | | | Parent | .60*** | 0.06 | .52*** | 0.09 |

Note. N: Neuroticism, A: Agreeableness, C: Conscientiousness, E: Extraversion, O: Openness. ** $p < .01$. CFI: Comparative Fit Index; RMSEA: Root Mean Square Error of Approximation.

* $p < .05$.
*** $p < .001$.

agreement during childhood and adolescence than less visible traits (i.e., neuroticism). Results also showed that by the age of

17, self-parent agreement in neuroticism had dropped to a non-significant and small level.

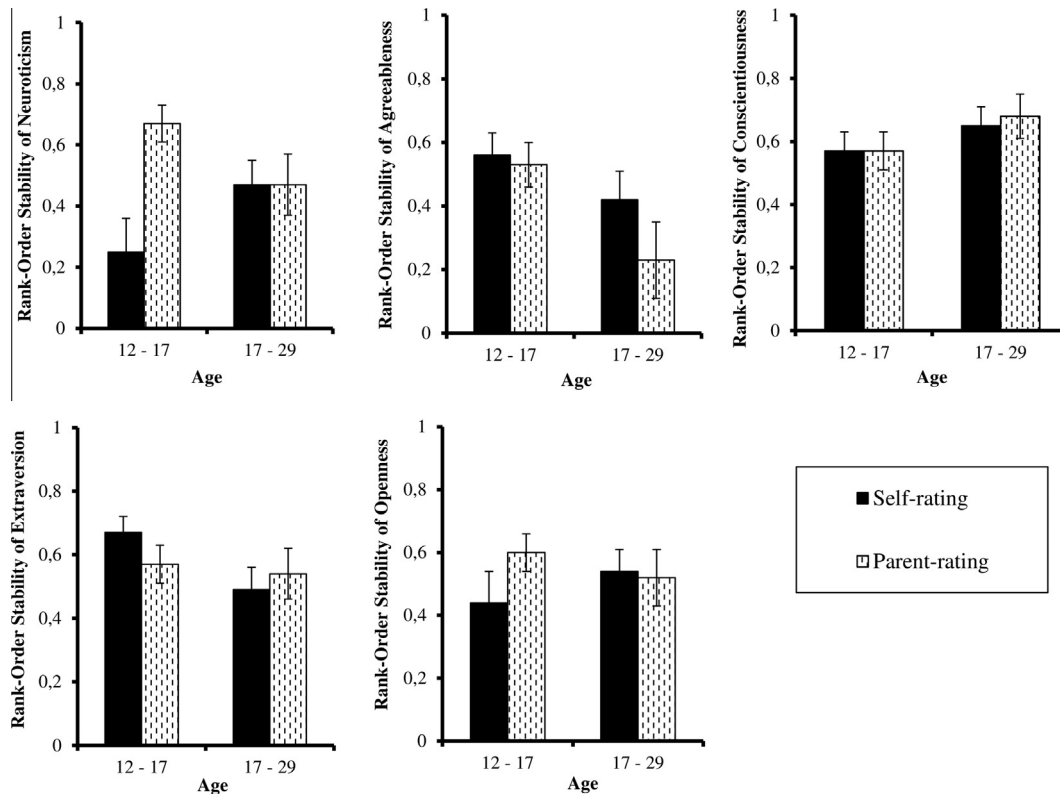


Fig. 2. Rank-order stability of the Big Five personality traits from age 12 to 29. Error bars represent ± 1 standard errors.

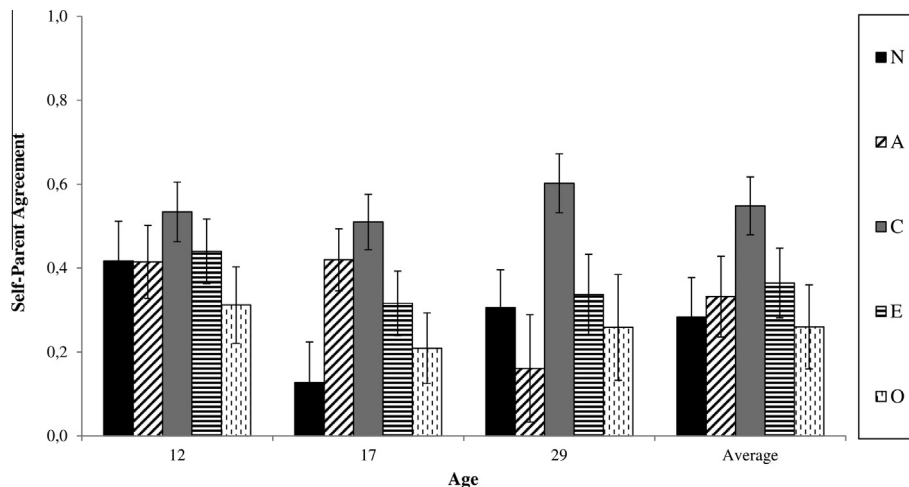


Fig. 3. Self-parent agreement in the Big Five personality traits from age 12 to 29. Average self-other agreements were computed using the Fisher r -to- z transformation. Error bars represent ± 1 standard errors.

Second, we examined the validity of the SOKA model under a developmental framework, that is: whether the *development* of more visible traits would also be more consistently judged than the *development* of less visible traits from multiple perspectives. Consistent with our hypothesis, we found the development of a more visible trait (i.e., extraversion), to be more consistently judged than of less visible traits (e.g., neuroticism), indicated by fewer differences in mean-level change and rank-order stability between children and their parents.

In summary, Study 1 provides valuable first insights, specifically that both the concurrent level and the development of children's personality depends on the judges, especially when it comes to less

visible traits such as neuroticism. In addition, results indicated the notion that personality maturation is a social construction and is observed differently by different judges. This long-term study, based on relatively long time intervals between measurements, has shown personality maturation in self-reports, and to an even greater degree, parent-reports.

However, due to the relatively small sample size of Study 1, replications are needed for the results we found, and a number of questions remain to be investigated. First, as suggested by the disruption hypothesis (Denissen et al., 2013; Soto & Tackett, 2015), personality maturation might not be without disturbance. Previous studies, zooming in on adolescence, have shown

temporary deviations from the maturity principle during adolescence, at least for some traits (Klimstra et al., 2009). To examine whether disruptions in the maturity principle can be found when using more fine-grained measures of personality and whether judges differ in their observations, Study 2 examined personality development as judged by multiple raters during adolescence with shorter consecutive measurements (i.e., annually).

In addition, Study 1 focused on self- and parent-ratings of children's personality development, leaving it unclear whether judges who are of a similar age to the adolescent targets would observe similar developmental patterns as the parents did. Adolescent siblings spend a lot of time with each other and encounter similar developmental tasks (Kim, McHale, Wayne Osgood, & Crouter, 2006). In Study 2 we examined whether siblings of similar age to the adolescents would confirm the maturity principle as observed by adolescents themselves.

4. Study 2

To closely test the disruption hypothesis across judges, Study 2 investigated personality developmental trajectories from multiple perspectives during adolescence. It was investigated annually, using an accelerated longitudinal design.

Are personality maturation and disruption during adolescence perceived differently across judges? First, we expected that at least some judges would observe deviation from the maturity principle during the “storm and stress” in adolescence. Second, we tested the notion that the maturity principle might somewhat be in the eye of the beholder. Given the increase in parent-child conflict during adolescence (Van der Giessen et al., 2014), we expected that parents would see less maturation of agreeableness, conscientiousness, and neuroticism than adolescents' themselves or their siblings.

Does the SOKA model apply to children and adolescence in a developmental framework? We first expected higher agreement between adolescents and their family members in more visible and behavioral oriented traits (e.g., extraversion and conscientiousness) than less visible traits (e.g., neuroticism). Second, we expected that the development (i.e., mean-level change and rank-order stability) of more visible traits (e.g., extraversion and conscientiousness) would be more consistently judged than the other personality traits (e.g., neuroticism).

4.1. Method

4.1.1. Participants and procedure

Participants were 576 Dutch adolescents (51% girls) who participated in the Family and Personality Research Project (Haselager & van Aken, 1999). A representative selection of 23 municipalities throughout the Netherlands provided lists of families with two adolescents between 11 and 16 years old. After a letter announcing the study, interviewers called families and invited them to participate, to which 50% of the contacted families agreed. The large majority of the respondents were of Dutch origin. In 4% of the families, parents reported that they were not born in the Netherlands (compared with 9% of the general Dutch population; Central Intelligence Agency, 2006).

At the start of the first measurement wave, the average ages for fathers and mothers were 43.9 years (ranging from 34.0 to 56.1 years old) and 41.7 years (ranging from 34.0 to 51.2 years old), respectively. Adolescents were classified into 5 cohorts by age: age 11.5 (ranging from 11.0 to 12.0 years old, $N = 99$), 12.5 (ranging from 12.0 to 13.0 years old, $N = 140$), 13.5 (ranging from 13.0 to 14.0 years old, $N = 114$), 14.5 (ranging from 14.0 to 15.0 years old, $N = 121$) and 15.5 (ranging from 15.0 to 16.0 years

old, $N = 102$). Participants were followed for three years, with annual measurements. In each measurement wave, trained experimenters visited the families at home and asked adolescents, their mothers, fathers, and adolescent siblings to judge the Big Five personality traits of the adolescents.

Most families participated throughout the entire course of the study: both at Wave 2 and Wave 3, 285 families (99%) provided data. Eighty-eight percent of the families provided complete data on the research variables at Wave 1, 99% at Wave 2, and 98% at Wave 3. In the same manner as in Study 1, we conducted a one-way ANOVA to compare complete cases (i.e., cases that showed no missingness for all research variables, $N = 490$) and cases with missingness (i.e., cases that showed missingness for at least one research variable, $N = 80$) on all research variables (at parcel level). Participants with complete cases showed: at Wave 1, higher mother-rated conscientiousness (Parcel 1 and 3), higher sibling-rated conscientiousness (Parcel 2), lower mother-rated neuroticism (Parcel 1); and at Wave 3, higher mother-rated conscientiousness (Parcel 1) and mother-rated extraversion (Parcel 2). For all other research variables, the 95% CI of complete cases overlapped with the 95% CI of cases with missingness. More detailed information can be found in [Table S2 in the online Supporting Information](#).

In addition, similar to Study 1, we conducted post-hoc power analyses with Monte Carlo simulation studies based on the estimated parameter value obtained from the present study. Regarding the mean-level change, results showed that for all Big Five personality traits, power was higher than .80 for investigating intercepts for each judge. Whether the slopes were significantly different from zero was tested using chi-square differences tests. However, calculating power for chi-square differences test is currently beyond the capability of the Monte Carlo simulation in Mplus. It might therefore be possible, that with a larger sample size, additional smaller but significant slopes for personality development from age 11.5 to 17.5 might be captured.

Regarding rank-order stability, results showed that for all Big Five personality traits, as rated by all the judges between all time intervals, power was higher than .80 for the investigation of rank-order stability. Regarding self-other agreement, results showed that for all Big Five personality traits at all measurement waves, power was higher than .80 for investigating significant self-other agreement.

4.1.2. Measures

Big Five personality traits. The Big Five personality traits (neuroticism, agreeableness, conscientiousness, extraversion, and openness to experience) of the participants were judged by the adolescents themselves, their mothers, fathers, and adolescent siblings using the Dutch adaptation (Gerris et al., 1998) of the 30 adjective Big Five personality traits (Goldberg, 1992). Sample items include “anxious” and “nervous” for neuroticism; “sympathetic” and “kind” for agreeableness; “careful” and “organized” for conscientiousness; “talkative” and “reserved” (reverse coded) for extraversion; and “imaginative” and “creative” for openness to experience. Personality was rated on a 7-point Likert scale (from 1 = *very untrue of this person*, to 7 = *very true of this person*).

In the same way as in Study 1, per Big Five trait the six items were parceled into three indicators using the Item-to-Construct Balancing approach (Little et al., 2002). The same parceling structure was applied to all the judges at all waves to ensure comparability. Cronbach's Alphas were satisfactory ([Table 1](#)).

4.1.3. Analytic strategy

Missing data handling and model fit. Missing data estimation and model fit indices were the same as reported for Study 1.

Mean-level change of the Big Five personality traits by the four judges. To investigate the mean-level change of personality,

multiple-group latent growth curve models were conducted for each personality trait separately in Mplus Version 7.11 (Muthén & Muthén, 2013). In each model, the five cohorts were treated as five groups and each cohort contributed three waves of data. Because each family had two participating adolescents, the “family id” was used as a cluster variable and the “analysis = complex” was applied to account for the dependency of observations.

The loadings of the intercepts for all ages (i.e., age 11.5–17.5) were fixed to 1. The loadings of the slopes at age 11.5 were fixed to 0 and at age 12.5 fixed to 0.1. All other loadings from age 13.5 to 17.5 were allowed to be freely estimated. The age – wave correspondence for each cohort can be found in Fig. 4. For example, age 12.5 corresponded to Wave 2 for Cohort 1, but corresponded to Wave 1 for Cohort 2. Measurement invariance specifications were included in the model, both across the three waves, as well as across the five cohorts in order to ensure that the psychometric meaning of the scale did not differ across waves and across cohorts.

Differences in intercepts and slopes between the four judges (i.e., self-rating, mother-rating, father-rating, and sibling-rating) were investigated using three dummy variables, representing the four judges. Self-rating was the reference group. Chi-square difference tests were applied to test whether intercepts and slopes were different between each pair of judges. The chi-square value cannot be analyzed in the regular manner when using the Robust Maximum Likelihood (MLR) estimator. Therefore, the procedure of Satorra-Bentler scaled chi-square difference test was applied for all model comparisons in Study 2 (Bryant & Satorra, 2012).

Rank-order stability of the Big Five personality traits by the four judges. We first conducted a multiple-group CFA with the specification of measurement invariance for each personality trait for each judge. In each model, the five cohorts were treated as five groups and each cohort contributed three waves of data. Rank-order stability of personality was investigated by adding correlations between the latent trait at Wave 1 and 2, as well as the correlations at Wave 2 and 3. The age – wave correspondence for each cohort can be found in Fig. 5. Again, the cluster function and the “analysis = complex” were applied to account for the dependency of observations. To test whether the rank-order stability during a

certain age period was different between each pair of judges, Satorra-Bentler scaled chi-square difference tests were applied to examine whether the correlation coefficients could be constrained to be equal across judges. Time lags were tested one by one.

Self-other agreement in the Big Five personality traits. We conducted CFAs with the specification of measurement invariance for each personality trait for each pair of judge separately. Self-other agreement in a certain personality trait refers to the correlation coefficients between the two latent factors (i.e., self-rating and other-rating).

4.2. Results

Means, standard deviations and inter-correlations of all manifest variables can be found in Table S3 in the online Supporting Information. The results concerning mean-level change of the Big Five personality traits from age 11.5 to 17.5, as judged by the adolescent themselves, their mothers, fathers and siblings are reported first, followed by the results of the rank-order stability of the Big Five personality traits. Finally, results of self-other agreements (i.e., self-mother, self-father, and self-sibling) in the Big Five personality traits are reported.

4.2.1. Mean-level development of the Big Five personality traits from multiple perspectives

Model fit indices and estimated mean-level change of the Big Five personality traits from age 11.5 to 17.5 can be found in Table 4 and Fig. 6.

Neuroticism. At age 11.5, parents perceived their children to be as neurotic as adolescents perceived themselves, but their siblings perceived them to be more neurotic. Detailed results of significant tests between self- and other-ratings for the intercepts can be found in Table 4.

Regarding the developmental trajectories from age 11.5 to 17.5, neuroticism remained stable for self- and sibling-ratings. By comparison, mothers judged their children to become less neurotic over time and fathers indicated a similar decreasing trend (mother: $\Delta\chi^2 = 5.47$, $\Delta df = 1$, $p = .019$; father: $\Delta\chi^2 = 2.05$, $\Delta df = 1$, $p = .152$).

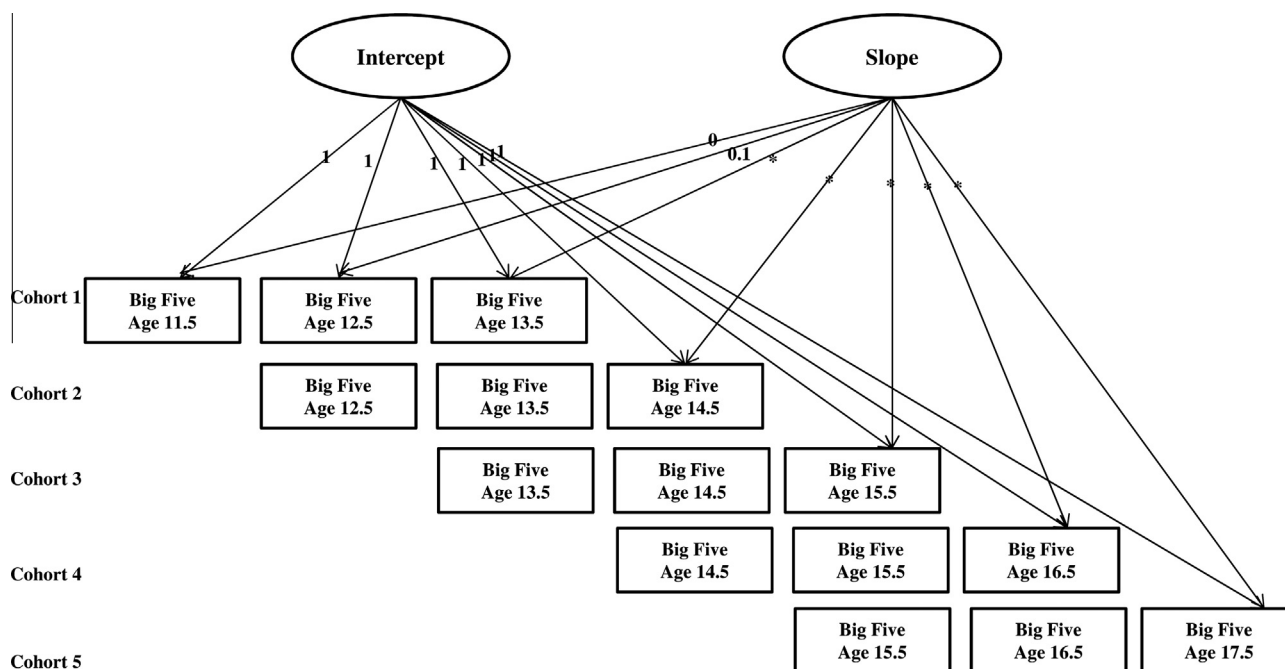


Fig. 4. Accelerated latent growth curve model for investigation of mean-level change of personality.

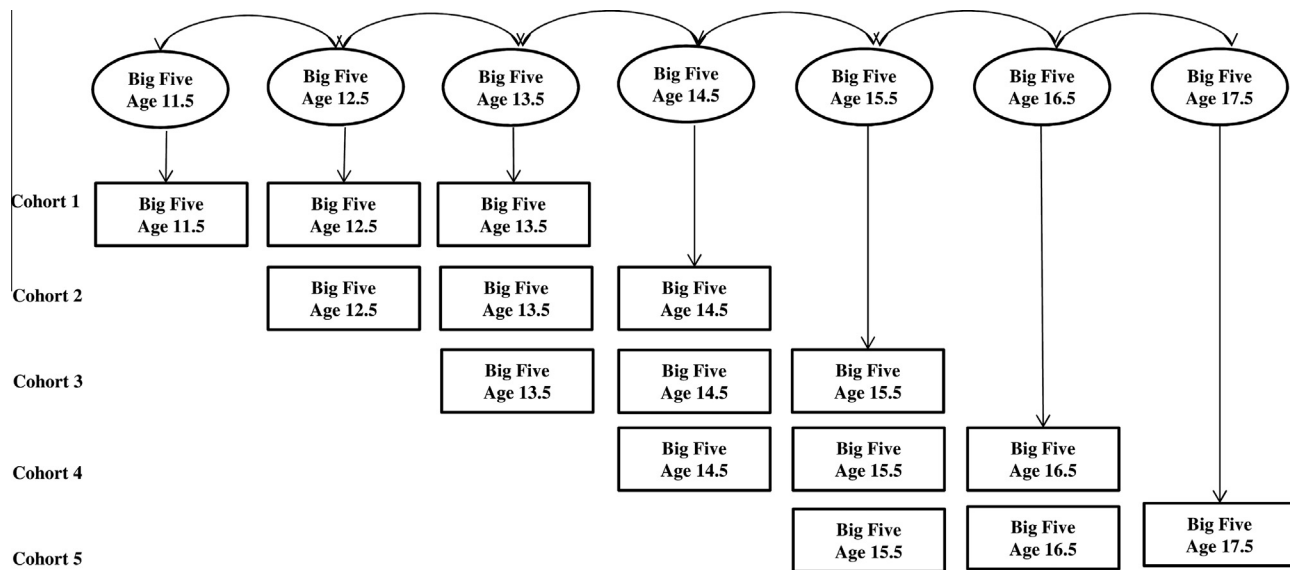


Fig. 5. Accelerated latent correlation model for investigation of rank-order stability of personality.

Agreeableness. At age 11.5 adolescents were seen as more agreeable by both parents, yet less agreeable by their siblings, as compared to adolescents self-views. From age 11.5 to 17.5, adolescents were seen by themselves and by their siblings as becoming increasingly agreeable over time, but were seen as becoming less agreeable over time by their parents (self: $\Delta\chi^2 = 10.88$, $\Delta df = 1$, $p < .001$; sibling: $\Delta\chi^2 = 24.11$, $\Delta df = 1$, $p < .001$; mother: $\Delta\chi^2 = 9.31$, $\Delta df = 1$, $p = .002$; father: $\Delta\chi^2 = 5.16$, $\Delta df = 1$, $p = .023$).

Conscientiousness. At age 11.5 adolescents judged themselves to be more conscientious than their mothers and siblings judged them, but not their fathers. From age 11.5 to 17.5, adolescents were seen by themselves and their siblings as becoming increasingly conscientious over time. However, they were seen by their parents as unchanged (self: $\Delta\chi^2 = 6.32$, $\Delta df = 1$, $p = .012$; sibling: $\Delta\chi^2 = 15.25$, $\Delta df = 1$, $p < .001$; mother: $\Delta\chi^2 = 0.37$, $\Delta df = 1$, $p = .543$; father: $\Delta\chi^2 = 0.26$, $\Delta df = 1$, $p = .610$).

Extraversion. Neither intercepts nor slopes differed between self-ratings and any of the other-ratings. The judgments of all judges indicated stability in the mean level of extraversion (mother vs. self: $b = -0.40$, $p = .842$, 95% CI [-4.39, 3.58], Cohen's $d = 0.03$; father vs. self: $b = -0.39$, $p = .844$, 95% CI [-4.22, 3.45], Cohen's $d = 0.03$; sibling vs. self: $b = -0.62$, $p = .793$, 95% CI [-5.26, 4.01], Cohen's $d = 0.05$).

Openness. Parents judged their children to be more open to experience than children judged themselves at age 11.5, while sibling-ratings did not differ from self-ratings. From ages 11.5 to 17.5, adolescents were seen by themselves and their siblings as becoming increasingly open to experience, but were seen as becoming less open to experience over time by their parents (self: $\Delta\chi^2 = 17.31$, $\Delta df = 1$, $p < .001$; sibling: $\Delta\chi^2 = 4.04$, $\Delta df = 1$, $p = .044$; mother: $\Delta\chi^2 = 5.22$, $\Delta df = 1$, $p = .022$; father: $\Delta\chi^2 = 8.85$, $\Delta df = 1$, $p = .003$).

4.2.2. Rank-order stability of the Big Five personality traits from multiple perspectives

Model fit indices and estimated rank-order stabilities of the Big Five personality traits from ages 11.5 to 17.5 can be found in Table 5 and Fig. 7.

Neuroticism. Compared to self-ratings, mother-ratings showed higher rank-order stabilities from ages 11.5 to 15.5. Father-ratings

showed higher rank-order stabilities from ages 12.5 to 15.5; and sibling-ratings showed lower rank-order stabilities from age 15.5 to 17.5. Detailed results of the significant tests can be found in Table 5.

Agreeableness. Compared to self-ratings, mother-ratings showed higher rank-order stabilities during ages 11.5–15.5. Father-ratings showed higher rank-order stabilities from age 11.5 to 13.5, from age 14.5 to 15.5, and from age 16.5 to 17.5. Finally, sibling-ratings showed lower rank-order stabilities from age 12.5 to 13.5, and from age 15.5 to 17.5.

Conscientiousness. Compared to self-ratings, the mother-ratings showed higher rank-order stability only from age 15.5 to 16.5. Father-ratings showed lower rank-order stabilities from age 16.5 to 17.5. Sibling-ratings showed lower rank-order stabilities from age 12.5 to 17.5.

Extraversion. Compared to self-ratings, the mother-ratings differed from self-ratings only during ages 14.5 to 15.5, with higher rank-order stability. Father-ratings differed only during ages 11.5–12.5, with higher rank-order stability. Sibling-ratings showed lower rank-order stabilities from age 15.5 to 16.5.

Openness. Compared to self-ratings, the mother-ratings showed higher rank-order stabilities from age 12.5 to 13.5 and from age 14.5 to 15.5. Father-ratings showed higher stabilities from age 12.5 to 13.5 but lower stabilities from age 16.5 to 17.5. Siblings-ratings showed lower rank-order stabilities from age 15.5 to 17.5.

4.2.3. Self-other agreement in the Big Five personality traits

As shown in Table 6 and Fig. 8, on average during adolescence self-other agreement was highest in conscientiousness and extraversion (both showed high agreement), lowest in neuroticism and agreeableness (both showed small to medium agreement). In addition, mothers showed higher agreement with adolescents themselves for all personality traits than fathers and siblings.

Also, as shown in Fig. 9, self-other agreement in neuroticism showed strong ups-and-downs during adolescence for all judges, especially for adolescents vs. siblings.

4.3. Discussion

Study 2 closely investigated the adolescents' personality development annually, as judged by adolescents themselves, their

Table 4
Study 2: Mean-level change of the Big Five personality traits by the four judges.

| Trait | χ^2 (df) | CFI | RMSEA | Judge | Means of intercept | Significant tests with the intercept of self-rating | | | Means of slope |
|-------|-----------------|-----|-------|---------|--------------------|---|----------------|-----------|----------------|
| | | | | | | p | 95% CI | Cohen's d | |
| N | 1136.17(365)*** | .92 | .07 | Self | 3.43 | – | – | – | 0.12 |
| | | | | Mother | 3.43 | .957 | [–0.17, 0.18] | 0.00 | –0.14** |
| | | | | Father | 3.47 | .643 | [–0.13, 0.21] | 0.06 | –0.14 |
| | | | | Sibling | 3.65 | .011 | [0.05, 0.39] | 0.27 | –0.15 |
| A | 1006.97(345)*** | .94 | .07 | Self | 5.47 | – | – | – | 0.16*** |
| | | | | Mother | 5.79 | .005 | [0.11, 0.53] | 0.47 | –0.12** |
| | | | | Father | 5.79 | .028 | [0.06, 0.58] | 0.40 | –0.20* |
| | | | | Sibling | 4.97 | <.001 | [–0.63, –0.37] | –0.63 | 0.26*** |
| C | 578.57(345)*** | .99 | .04 | Self | 4.15 | – | – | – | 0.05* |
| | | | | Mother | 4.00 | .031 | [–0.29, –0.01] | –0.21 | –0.01 |
| | | | | Father | 4.10 | .516 | [–0.19, 0.09] | 0.07 | –0.00 |
| | | | | Sibling | 3.74 | <.001 | [–0.56, –0.26] | –0.54 | 0.09*** |
| E | 494.04(299)*** | .98 | .04 | Self | 4.81 | – | – | – | 0.40 |
| | | | | Mother | 4.83 | .954 | [–0.35, 0.37] | 0.01 | –0.00 |
| | | | | Father | 4.84 | .883 | [–0.33, 0.38] | 0.02 | 0.02 |
| | | | | Sibling | 4.85 | .892 | [–0.47, 0.54] | 0.03 | –0.22 |
| O | 962.12 (345)*** | .94 | .06 | Self | 4.99 | – | – | – | 0.43*** |
| | | | | Mother | 5.21 | .015 | [0.04, 0.41] | 0.32 | –0.25* |
| | | | | Father | 5.34 | <.001 | [0.20, 0.51] | 0.56 | –0.27** |
| | | | | Sibling | 4.93 | .393 | [–0.18, 0.07] | 0.09 | 0.22* |

Note. N: Neuroticism, A: Agreeableness, C: Conscientiousness, E: Extraversion, O: Openness. CFI: Comparative Fit Index; RMSEA: Root Mean Square Error of Approximation.

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

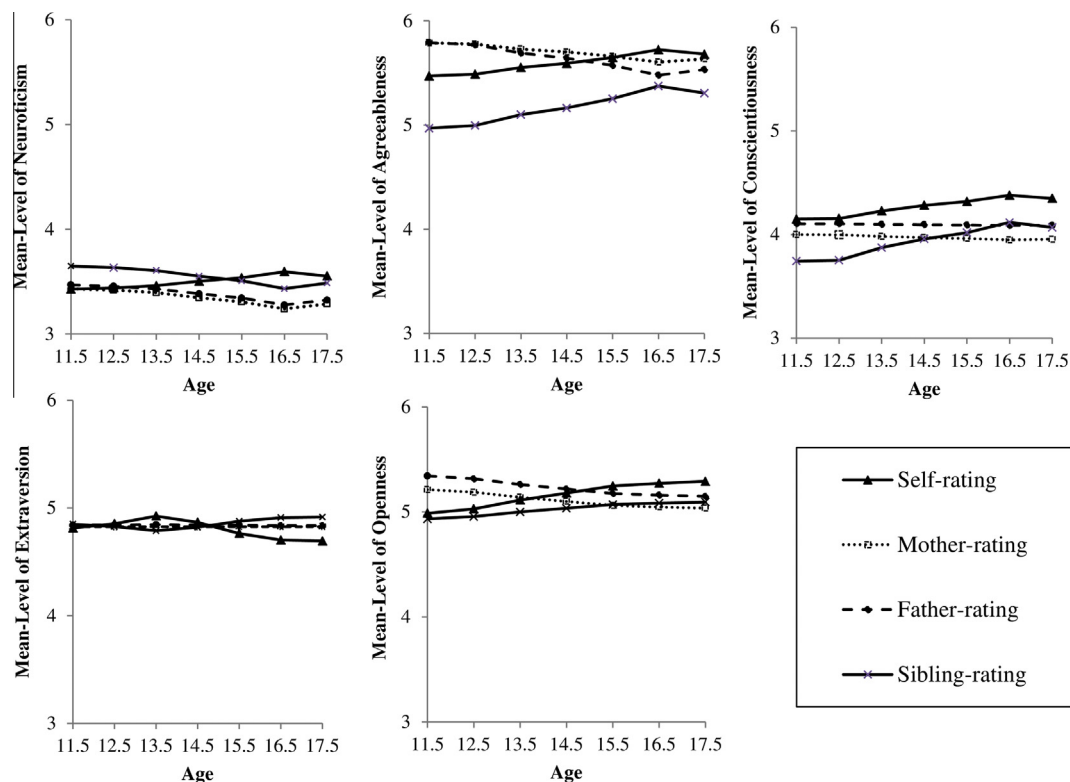


Fig. 6. Mean-level change of the Big Five personality traits from age 12 to 18.

mothers, fathers, and siblings. We discussed the results below in accordance with our three research questions.

4.3.1. Are personality maturation and disruption during adolescence perceived differently across judges?

First, our results confirmed the disruption hypothesis (Denissen et al., 2013; Soto & Tackett, 2015) by revealing dips in personality

maturation, observed by at least some of the judges in some of the traits. It was only a partial confirmation, however, because only mothers and fathers observed the decreases in agreeableness during adolescence that are associated with the disruption hypothesis.

Second, our results supported the notion that the maturity principle is a social construction to some degree, by showing different

Table 5
Study 2: Rank-order stability of the Big Five personality traits by the four judges.

| Trait | Judge | χ^2 (df) | CFI | RMSEA | Significant tests with the rank-order stability of self-rating: TRD (Δdf) | | | | | |
|-------|---------|----------------|------|-------|---|---------------|---------------|---------------|---------------|---------------|
| | | | | | Age 11.5–12.5 | Age 12.5–13.5 | Age 13.5–14.5 | Age 14.5–15.5 | Age 15.5–16.5 | Age 16.5–17.5 |
| N | Self | 359.77(234)*** | .93 | .07 | – | – | – | – | – | – |
| | Mother | 392.86(234)*** | .94 | .08 | 3.94(1)* | 22.62(1)*** | 12.81(1)*** | 4.10(1)* | Non-sig. | Non-sig. |
| | Father | 378.06(234)*** | .94 | .07 | Non-sig. | 6.21(1)* | 3.97(1)* | 14.91(1)*** | Non-sig. | Non-sig. |
| | Sibling | 384.66(234)*** | .91 | .07 | Non-sig. | Non-sig. | Non-sig. | Non-sig. | 7.57(1)** | 20.42(1)*** |
| A | Self | 382.35(214)*** | .90 | .08 | – | – | – | – | – | – |
| | Mother | 345.67(214)*** | .96 | .07 | 7.35(1)** | 9.53(1)** | 9.22(1)** | 12.78(1)*** | Non-sig. | Non-sig. |
| | Father | 309.29(214)*** | .97 | .06 | 5.19(1)* | 19.98(1)*** | Non-sig. | 14.14(1)*** | Non-sig. | 4.66(1)* |
| | Sibling | 331.86(214)*** | .96 | .07 | Non-sig. | 6.31(1)** | Non-sig. | Non-sig. | 7.42(1)** | 7.82(1)** |
| C | Self | 379.14(213)*** | .95 | .08 | – | – | – | – | – | – |
| | Mother | 304.60(213)*** | .98 | .06 | Non-sig. | Non-sig. | Non-sig. | Non-sig. | 5.21(1)* | Non-sig. |
| | Father | 218.13(213) | 1.00 | .01 | Non-sig. | Non-sig. | Non-sig. | Non-sig. | Non-sig. | 5.69(1)* |
| | Sibling | 328.13(213)*** | .96 | .07 | Non-sig. | 9.50(1)** | 14.74(1)*** | 18.23(1)*** | 4.79(1)* | 6.17(1)* |
| E | Self | 352.14(214)*** | .94 | .08 | – | – | – | – | – | – |
| | Mother | 280.53(214)** | .98 | .05 | Non-sig. | Non-sig. | Non-sig. | 4.95(1)* | Non-sig. | Non-sig. |
| | Father | 240.13(214) | .99 | .03 | 6.08(1)* | Non-sig. | Non-sig. | Non-sig. | Non-sig. | Non-sig. |
| | Sibling | 276.99(214)** | .97 | .05 | Non-sig. | Non-sig. | Non-sig. | Non-sig. | 5.69(1)* | Non-sig. |
| O | Self | 356.79(209)*** | .91 | .08 | – | – | – | – | – | – |
| | Mother | 362.46(209)*** | .95 | .08 | Non-sig. | 14.23(1)*** | Non-sig. | 6.76(1)** | Non-sig. | Non-sig. |
| | Father | 336.89(209)*** | .96 | .07 | Non-sig. | 14.19(1)*** | Non-sig. | Non-sig. | Non-sig. | 10.85(1)** |
| | Sibling | 272.68(209)** | .96 | .05 | Non-sig. | Non-sig. | Non-sig. | 13.63(1)*** | 8.40(1)** | 5.74(1)* |

Note. N: Neuroticism, A: Agreeableness, C: Conscientiousness, E: Extraversion, O: Openness. Non-sig. indicated $p \geq .05$. CFI: Comparative Fit Index; RMSEA: Root Mean Square Error of Approximation. Chi-square differences tests were applied by using the Satorra-Bentler Scaled Chi-Square. $TRD = (T_0 * C_0 - T_1 * C_1) / cd$, $cd = (d_0 * c_0 - d_1 * c_1) / (d_0 - d_1)$.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

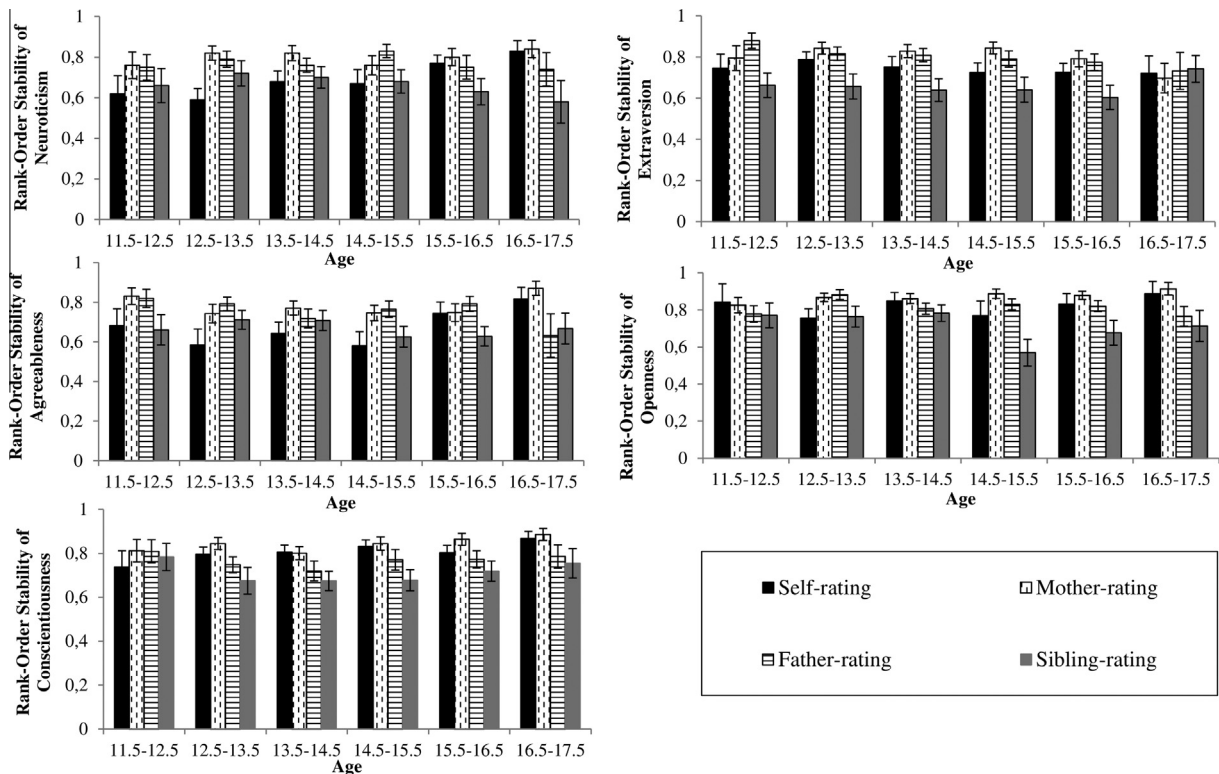


Fig. 7. Rank-order stability of the Big Five personality traits from age 12 to 18. Error bars represent ± 1 standard errors.

degrees of or aspects of the maturity principle for different judges. In our study, adolescents saw themselves as increasingly agreeable and conscientious, whereas they were seen by their parents as decreasingly agreeable and neurotic.

Referring to the items of the personality scale used in the present study, these results indicated that parents see maturation

of adolescents' personality as being less stressed, insecure and worried than before. But parents also saw adolescents become increasingly reckless, argumentative, and irritable. Adolescents' views were different. The adolescents regarded themselves as being just as emotionally vulnerable and stressed as they were before. However, they see maturation in the way they regulate

Table 6

Study 2: Self-other agreement in the Big Five personality traits.

| Trait | Self-other agreement | χ^2 (df) | CFI | RMSEA | Age 11.5 | Age 12.5 | Age 13.5 | Age 14.5 | Age 15.5 | Age 16.5 | Age 17.5 |
|-------|----------------------|-----------------------------|------|-------|----------|----------|----------|----------|----------|----------|----------|
| N | Self-mother | 956.46(653) ^{***} | 0.95 | 0.06 | .24 | .33 | .36 | .28 | .28 | .46 | .47 |
| | Self-father | 1157.12(698) ^{***} | 0.92 | 0.08 | .12 | .08 | .30 | .21 | .24 | .22 | .31 |
| | Self-sibling | 1097.01(698) ^{***} | 0.92 | 0.07 | .04 | .40 | .26 | .11 | .20 | .23 | .36 |
| A | Self-mother | 1231.66(698) ^{***} | 0.92 | 0.08 | .17 | .22 | .25 | .31 | .28 | .42 | .55 |
| | Self-father | 1124.19(698) ^{***} | 0.93 | 0.07 | .06 | .25 | .29 | .20 | .24 | .27 | .15 |
| | Self-sibling | 1262.98(697) ^{***} | 0.91 | 0.08 | .11 | .17 | .17 | .25 | .32 | .35 | .21 |
| C | Self-mother | 989.70(698) ^{***} | 0.97 | 0.06 | .53 | .60 | .59 | .52 | .65 | .55 | .63 |
| | Self-father | 990.28(698) ^{***} | 0.97 | 0.06 | .55 | .58 | .55 | .43 | .57 | .48 | .55 |
| | Self-sibling | 1053.84(698) ^{***} | 0.95 | 0.07 | .36 | .50 | .47 | .44 | .50 | .54 | .52 |
| E | Self-mother | 1003.18(698) ^{***} | 0.96 | 0.06 | .59 | .57 | .47 | .41 | .54 | .50 | .57 |
| | Self-father | 894.58(698) ^{***} | 0.97 | 0.05 | .53 | .39 | .45 | .38 | .43 | .34 | .32 |
| | Self-sibling | 951.12(698) ^{***} | 0.96 | 0.06 | .51 | .53 | .42 | .34 | .38 | .30 | .30 |
| O | Self-mother | 921.42(653) ^{***} | 0.96 | 0.06 | .52 | .47 | .48 | .47 | .38 | .39 | .28 |
| | Self-father | 915.96(653) ^{***} | 0.96 | 0.06 | .44 | .47 | .37 | .41 | .42 | .25 | .18 |
| | Self-sibling | 876.25(622) ^{***} | 0.94 | 0.06 | .28 | .37 | .51 | .40 | .36 | .53 | .48 |

Note. N: Neuroticism, A: Agreeableness, C: Conscientiousness, E: Extraversion, O: Openness. * $p < .05$, ** $p < .01$. CFI: Comparative Fit Index; RMSEA: Root Mean Square Error of Approximation.

^{***} $p < .001$.

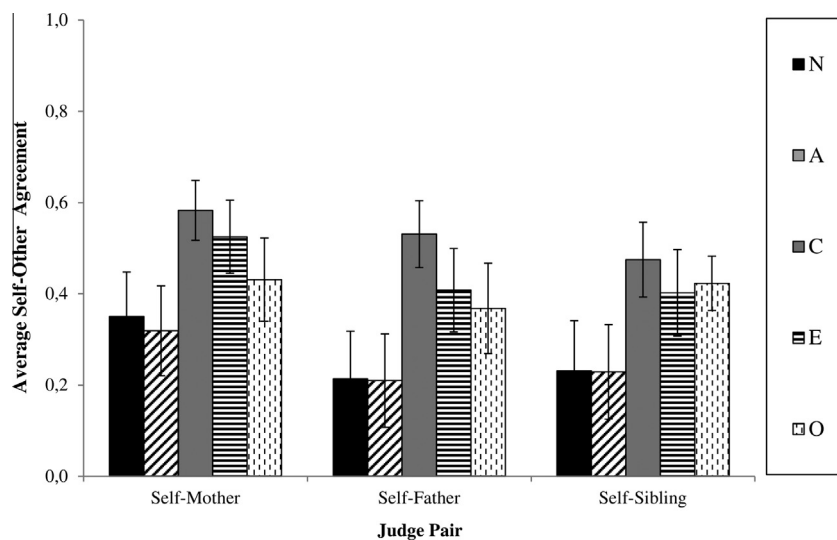


Fig. 8. Average self-other agreement in the Big Five personality traits from age 12 to 18. Average self-other agreements were computed using the Fisher r -to- z transformation. Error bars represent ± 1 standard errors.

their interpersonal interactions and daily tasks, shown through increases in the two self-regulatory traits: agreeableness and conscientiousness (Soto & Tackett, 2015).

Regarding the siblings' view of adolescents' personality maturation, the results confirmed our expectation that siblings observe similar personality maturation as the adolescents themselves. Unexpectedly, concerning the intercepts, we found siblings' judgments at age 11.5 to be more negative than adolescents' self-views⁴ in conscientiousness, neuroticism and agreeableness. These more negative views of the siblings on the maturity of personality traits might be related to the competitiveness between siblings for parental resources (e.g., attention, love, money) during childhood, which was shown to decline before early adolescence (Kim et al., 2006).

⁴ Siblings' judgments on these three traits at age 11.5 were also more negative than parents' judgments based on additional analyses. Detailed results can be retrieved from the first author upon request.

4.3.2. Does the SOKA model apply to children and adolescents in a developmental framework?

Our results confirmed the SOKA Model (Vazire, 2010) in an adolescent sample and extended its validity in the developmental context. Specifically, we first found higher self-other agreement during adolescence for more visible and behavioral-oriented traits (e.g., extraversion and conscientiousness) than for less visible traits (e.g., neuroticism).

Second, we found that more visible traits were not only concurrently, but also longitudinally, more consistently judged than less visible traits among an adolescent's family members, indicated by fewer differences in the mean-level change and rank-order stability between judges. In addition, adolescents' judgments again showed lower rank-order stabilities than their parents', which may be due to their highly fluctuating emotional status and their less committed identity. Adolescents' judgments were less stable than parents' judgments not only when they were judging themselves, but also when they were judging their siblings.

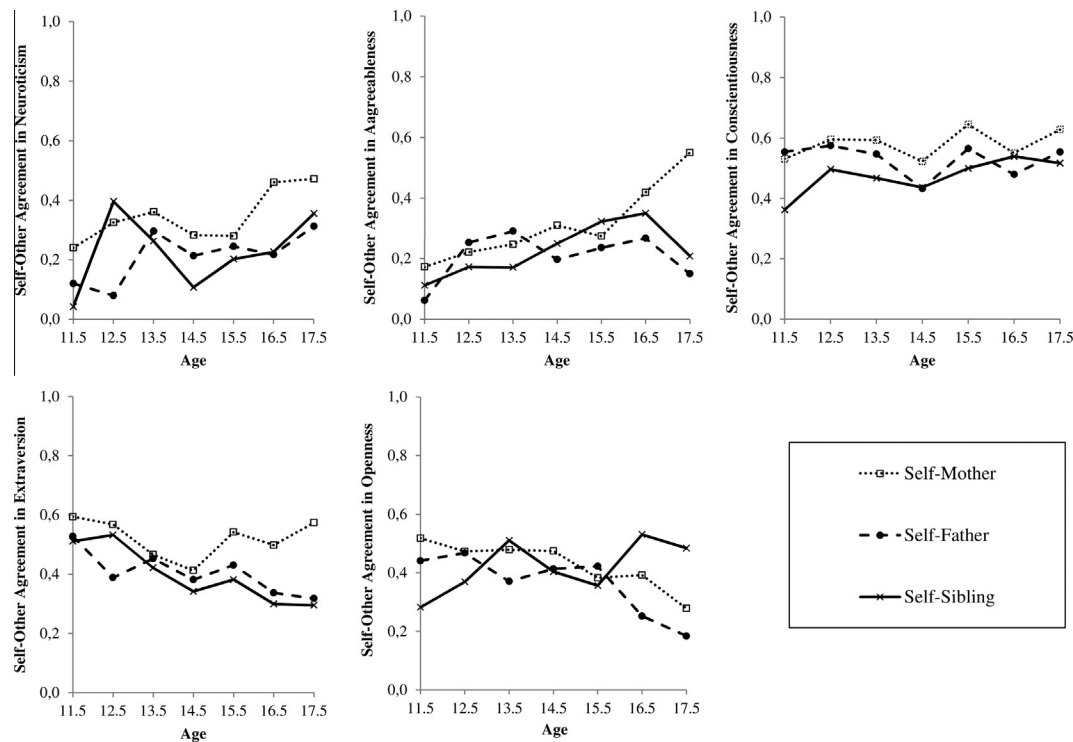


Fig. 9. Self-other agreement in the Big Five personality traits from age 12 to 18.

5. General discussion

In two longitudinal studies, we consistently found mean-level change and rank-order stability of personality to depend on the judge in meaningful ways. Looking at long-term development (Study 1), personality development from late childhood to young adulthood through both self- and parent- reports confirmed the maturity principle. However, when zooming in on adolescence (Study 2), maturation was disrupted during adolescence. In addition, we confirmed the SOKA Model in the children and adolescents sample and expanded its validity into the developmental context. In the following, we discuss the results in more detail.

5.1. Are personality maturation and disruption from childhood to young adulthood perceived differently across judges?

From childhood to young adulthood, our findings supported the maturity principle (Roberts et al., 2006) by showing that both children themselves, and (to a greater degree) their parents' judgments showed personality maturation over time. That is, parent-ratings showed a similar amount of increase in agreeableness, but a greater increase in conscientiousness than the self-ratings. Moreover, parent-ratings perceived decreases in neuroticism, whereas the children reported that they were just as emotionally stressed and vulnerable as they were before. These results indicate that in the long run, parents see their children's personality development in a more positive light than the children themselves.

However, personality maturation is not without disruption, as suggested by the disruption hypothesis (Denissen et al., 2013; Soto & Tackett, 2015). Indeed, when taking a closer look at personality development in adolescence, we found that adolescents were seen by their mothers and fathers as becoming less agreeable (and also less open to experience) over time.

Our results supported the suggestion from previous studies (Watson & Humrichouse, 2006) that the maturity principle might

(at least to some degree) be a social construction, meaning that different judges may vary in the degree to, and the traits in which, they observe the maturity principle and the disruption of it. Parents saw maturation in the way their adolescent children experience emotions (decreases in neuroticism). Adolescents' reports, however, showed that they were just as emotionally vulnerable and stressed as they were before. Instead, adolescents (and their siblings) saw maturation in their self-regulatory traits (increases in agreeableness and conscientiousness), whereas their parents did not.

How can these differences be explained? One speculative explanation is that as adolescents' strive for autonomy and independence from their parents, their inner feelings and sufferings are no longer very accessible to their parents (Van der Giessen et al., 2014), making it hard for parents to sympathize with changes in emotional traits such as neuroticism (Vazire, 2010). In addition, adolescents go through a number of biological and socio-emotional transitions, and it takes some time before adolescents' increasing self-regulatory capacities become sufficient for properly regulating a majority of their emotions and social tasks (Denissen et al., 2013). Therefore, the increases of adolescent's self-regulatory traits might not be immediately visible to their parents.

It should be noted that we are not assuming that one judgment is more "correct" than the other. Previous studies have shown that each perspective contains unique information, and no perspective is better than the rest. Rather, which perspective is pragmatically more useful is highly dependent on the specific outcome that is being targeted (Connelly & Ones, 2010; Vazire, 2010). The discrepancy we found between parents and adolescents, however, might have implications for understanding and promoting parent-adolescent relationship quality, as self-verifying feedback from close others boosts feelings of being understood and relationship quality (Gordon & Chen, 2015; Human & Biesanz, 2013; Swann, De La Ronde, & Hixon, 1994). Failure of parents to sympathize with their adolescents' inner stress and vulnerability might limit their ability to provide the necessary support.

A recent study focusing on a younger and shorter age period (from 10.70 to 13.70 years old) found several similar results – neuroticism declined in parent-ratings, but remained stable in self-reports; agreeableness decreased and conscientiousness remained stable in parent-ratings (Göllner et al., in press). However, different from our results, decreases were found in the self-ratings of conscientiousness (rather than increases) and agreeableness (rather than remaining stable). Knowledge from both studies suggest the disruption in personality maturation to be in the eyes of the beholder, in addition to being a temporary phenomenon. Looking at longer-term development, both children themselves and (to a greater degree) parents perceived personality maturation. More research with additional judges and age groups is needed to replicate and further illuminate this matter.

5.2. Does the SOKA model apply to children and adolescents in a developmental framework?

Our data have interesting implications for the generalizability of the SOKA Model (Vazire, 2010) – which was based on cross-sectional studies on adults – to children and adolescents. First, consistent with our hypotheses, in both studies we found higher self-other agreement in more visible and behavioral-oriented traits (e.g., conscientiousness and extraversion) than in less visible traits (e.g., neuroticism) between children and their family members. Especially at age 17, self-parent agreement had dropped to non-significant and small level.

An unexpected but potentially interesting finding was that although self-parent agreement in agreeableness was significant and at a medium level at age 12 and 17, this agreement dropped to non-significant and was only at a small level at age 29. One explanation might be that the interaction with parents at this later age becomes a smaller part of individuals' social lives, than compared to childhood and adolescence. A person might behave prosocially with their parents (and other family members), but less so with friends, colleagues and strangers, or vice versa. Certainly this speculation needs future studies to verify.

Second, results confirmed and expanded the predictions from the SOKA Model to personality development in childhood and adolescence, by showing fewer inter-judge differences in the development (mean-level change and rank-order stability) of more visible traits (i.e., extraversion and conscientious) than the development of less visible and more evaluative traits (e.g., neuroticism and openness).

These results can be explained using the framework of the realistic accuracy model (Funder, 1995), which suggests that a personality judgment contains four steps: relevant information exists for that trait (relevance), is available to the judge (availability), is noticed by the judge (detection), and is interpreted correctly (utilization). More visible traits (e.g. extraversion and conscientiousness) ease steps such as relevance and availability (e.g., extraverts' exaggerated facial expressions and loud voice, and a conscientious person's tidy office). When such trait-relevant cues are not sufficiently available (such as for less visible traits), judgments are more influenced by various heuristics, such as the current relationship quality (Watson et al., 2000), leading to less consistency in judgments.

Third, although it was not the focus of our study, an additional interesting finding was the fluctuating characteristic of adolescents' personality judgments. Both studies showed that children's personality judgments showed lower rank-order stability during adolescence than parents' judgments. In Study 1 self-views were less stable than the parent views of neuroticism during adolescence. Study 2 confirmed this finding and further showed that self-views on other traits (especially openness and agreeableness)

were also less stable than parents' views. Furthermore, siblings' views were even less stable than self-views.

One possibility is that adolescents' emotional fluctuations together with their less committed identity made adolescents' self-views less stable than their parents'. In addition, adolescents' emotional fluctuations might not only be reflected in lower rank-order stabilities (e.g., compared to their parents') in judging their own personality, but also in the impressions they form for others (i.e., their siblings), suggesting a fluctuating characteristic of adolescent judges. That is, adolescents' judgments of their own and their siblings' personality might fluctuate more with their current mood or state self-esteem than their parents' judgments.

5.3. Limitations and future directions

The present article is the first to shed light on the developmental trajectories of the Big Five personality traits from childhood to young adulthood as judged from multiple perspectives in two longitudinal studies. However, it also contains some limitations. First, both studies focused on judgments of personality development by children themselves and their family members. Future studies might be interested in examining differences between other important judges in adolescence (e.g. friends and romantic partners) and extending results to other life phases (e.g., mid- and late-adulthood). Second, the present study did not establish the external validity of personality judgment by different judges, such as their predictive validity for future outcomes. However, previous studies have shown that there is no best perspective, given that each perspective provides unique information (Connelly & Ones, 2010; Vazire & Mehl, 2008). Future studies might want to examine whose personality judgments of specific traits outperform the others' in predicting different types of important developmental outcomes. Third, the sample sizes of both of our studies were modest. It is crucial for future studies with larger sample sizes to replicate our findings. Finally, the present study was unable to directly address the dynamic transaction between relationship quality and personality judgment – such as the correlated changes between self-other relationship quality and self-other agreement on personality development – which would be an interesting question for future studies.

6. Conclusion

Results from our two longitudinal studies confirm and extend the self-other knowledge asymmetry model to childhood and adolescence in a developmental framework. Results indicated that personality maturation and the disruption of it is to some degree a social construction, since judges saw different degrees and aspects of the maturity principle. Specifically, in the longer term – from childhood to young adulthood – both children and (to a greater degree) parents saw personality maturation, although it was not without disruption. Zooming in on the “storm and stress” phase of adolescence, we found parent reports of personality development to confirm the disruption hypothesis by showing decreases in agreeableness (and also openness). In addition, parents' views showed maturation in the way adolescents experience emotions but not in the way they cooperate with others (i.e., decreases in neuroticism and agreeableness). However, adolescents' self-views did not show decreasing emotional vulnerability, but showed maturation in their self-regulatory traits (i.e., increases in agreeableness and conscientiousness). Understanding the different views in personality development between adolescents and their family members might have theoretical and practical implications for improving children's and adolescent's feeling of being understood and their social relationships.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jrp.2016.03.004>.

References

- Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, 55, 469–480.
- Asendorpf, J. B., Denissen, J. J. A., & van Aken, M. A. G. (2008). Inhibited and aggressive preschool children at 23 years of age: Personality and social transitions into adulthood. *Developmental Psychology*, 44(4), 997–1011. <http://dx.doi.org/10.1037/0012-1649.44.4.997>.
- 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(12), 2530–2540. <http://dx.doi.org/10.1177/0956797613498396>.
- Branje, S. J. T., van Lieshout, C. F. M., & Gerris, J. R. M. (2007). Big Five personality development in adolescence and adulthood. *European Journal of Personality*, 21(1), 45–62. <http://dx.doi.org/10.1002/per.596>.
- Bryant, F. B., & Satorra, A. (2012). Principles and practice of scaled difference chi-square testing. *Structural Equation Modeling: A Multidisciplinary Journal*, 19(3), 372–398. <http://dx.doi.org/10.1080/10705511.2012.687671>.
- Carlson, E. N., & Furr, R. M. (2009). Evidence of differential meta-accuracy: People understand the different impressions they make. *Psychological Science*, 20(8), 1033–1039. <http://dx.doi.org/10.1111/j.1467-9280.2009.02409.x>.
- Carlson, E. N., Vazire, S., & Furr, R. M. (2011). Meta-insight: Do people really know how others see them? *Journal of Personality and Social Psychology*, 101(4), 831–846. <http://dx.doi.org/10.1037/a0024297>.
- Casey, B. J., Jones, R. M., Levita, L., Libby, V., Pattwell, S. S., Ruberry, E. J., ... Somerville, L. H. (2010). The storm and stress of adolescence: Insights from human imaging and mouse genetics. *Developmental Psychobiology*, 52(3), 225–235. <http://dx.doi.org/10.1002/dev.20447>.
- Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, 136(6), 1092–1122. <http://dx.doi.org/10.1037/a0021212>.
- Denissen, J. J. A. (2014). Editorial: A roadmap for further progress in research on personality development. *European Journal of Personality*, 215(28), 213–215. <http://dx.doi.org/10.1002/per.1967>.
- Denissen, J. J. A., Asendorpf, J. B., & van Aken, M. A. G. (2008). Childhood personality predicts long-term trajectories of shyness and aggressiveness in the context of demographic transitions in emerging adulthood. *Journal of Personality*, 76(1), 67–99. <http://dx.doi.org/10.1111/j.1467-6494.2007.00480.x>.
- 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(4), 255–260. <http://dx.doi.org/10.1111/cdep.12050>.
- Denissen, J. J. A., van Aken, M. A. G., & Roberts, B. W. (2011). Personality development across the life span. In T. Chamorro-Premuzic, A. Furnham, & S. von Stumm (Eds.), *Handbook of individual differences* (pp. 77–100). Oxford: Wiley-Blackwell.
- Durbin, C. E., & Wilson, S. (2012). Convergent validity of and bias in maternal reports of child emotion. *Psychological Assessment*, 24(3), 647–660. <http://dx.doi.org/10.1037/a0026607>.
- Ferguson, C. J. (2010). A meta-analysis of normal and disordered personality across the life span. *Journal of Personality and Social Psychology*, 98(4), 659–670. <http://dx.doi.org/10.1037/a0018770>.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, 102(4), 652–670. <http://dx.doi.org/10.1037/0033-295X.102.4.652>.
- Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology*, 55(1), 149–158. <http://dx.doi.org/10.1037/0022-3514.55.1.149>.
- Funder, D. C., & Sneed, C. D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology*, 64(3), 479–490.
- Gerris, J. R. M., Houtmans, M. J. M., Kwaaitaal-Roosen, E. M. G., Schipper, J. C., Vermulst, A. A., & Janssens, J. M. A. M. (1998). *Parents, adolescents, and young adults in Dutch families: A longitudinal study*. Nijmegen, the Netherlands: University of Nijmegen, Institute of Family Studies.
- Goldberg, L. R. (1992). The development of markers of the Big-Five factor structure. *Psychological Assessment*, 4, 26–42. <http://dx.doi.org/10.1037/1040-3590.4.1.26>.
- Göllner, R., Roberts, B. W., Damian, R. I., Lüdtke, O., Jonkmann, K., & Trautwein, U. (in press). Whose "storm and stress" is it? Parent and child reports of personality development in the transition to early adolescence. *Journal of Personality* (in press). Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/jopy.12246/abstract>.
- Gordon, A. M., & Chen, S. (2015). Do you get where I'm coming from? Perceived understanding buffers against the negative impact of conflict on relationship satisfaction. *Journal of Personality and Social Psychology*. <http://dx.doi.org/10.1037/pspi0000039>. Advance online publication.
- Haselager, G. J. T., & van Aken, M. A. G. (1999). *Family and personality: First measurement wave*. Nijmegen, the Netherlands: University of Nijmegen.
- Human, L. J., & Biesanz, J. C. (2013). Targeting the good target: An integrative review of the characteristics and consequences of being accurately perceived. *Personality and Social Psychology Review*, 17(3), 248–272. <http://dx.doi.org/10.1177/1088868313495593>.
- Jenkins, J., & Dunn, J. (2009). Siblings within families: Levels of analysis and patterns of influence. *New Directions for Child and Adolescent Development*, 126, 79–93. <http://dx.doi.org/10.1002/cd.258>.
- John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The big five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, 61(4), 521–551. <http://dx.doi.org/10.1111/j.1467-6494.1993.tb00781.x>.
- Kim, J.-Y., McHale, S. M., Wayne Osgood, D., & Crouter, A. C. (2006). Longitudinal course and family correlates of sibling relationships from childhood through adolescence. *Child Development*, 77(6), 1746–1761. <http://dx.doi.org/10.1111/j.1467-8624.2006.00971.x>.
- Klimstra, T. A., Hale, W. W., Raaijmakers, Q. A. W., Branje, S. J. T., & Meeus, W. H. (2009). Maturation of personality in adolescence. *Journal of Personality and Social Psychology*, 96(4), 898–912. <http://dx.doi.org/10.1037/a0014746>.
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, 9(2), 151–173. http://dx.doi.org/10.1207/S15328007SEM0902_1.
- Lucas, R. E., & Donnellan, M. B. (2011). Personality development across the life span: Longitudinal analyses with a national sample from Germany. *Journal of Personality and Social Psychology*, 101(4), 847–861. <http://dx.doi.org/10.1037/a0024298>.
- Marsh, H. W., Hau, K., & Grayson, D. M. (2005). Goodness of fit in Structural Equation Models. In: A. Maydeu-Olivares, & J. McArdle (Eds.), *Mahwah*.
- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo Study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(4), 599–620. http://dx.doi.org/10.1207/S15328007SEM0904_8.
- Muthén, L. K., & Muthén, B. O. (2013). *Mplus 7.11*. Los Angeles, CA: Muthén & Muthén.
- Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit [Language and personality structure: Validity of the Five Factor Model of personality]*. Regensburg, Germany: Roderer.
- Roberts, B. W., & DeVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, 126(1), 3–25. <http://dx.doi.org/10.1037/0033-2909.126.1.3>.
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, 132(1), 1–25. <http://dx.doi.org/10.1037/0033-2909.132.1.1>.
- Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 375–398). New York, NY: Guilford Press.
- Soto, C. J., & Tackett, J. L. (2015). Personality traits in childhood and adolescence: Structure, development, and outcomes. *Current Directions in Psychological Science*, 24(5), 358–362. <http://dx.doi.org/10.1177/0963721415589345>.
- Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology*, 101(4), 862–882. <http://dx.doi.org/10.1037/a0024950>.
- Swann, J. W. B., De La Ronde, C., & Hixon, J. G. (1994). Authenticity and positivity strivings in marriage and courtship. *Journal of Personality and Social Psychology*, 66(5), 857–869. <http://dx.doi.org/10.1037/0022-3514.66.5.857>.
- Tackett, J. L. (2011). Parent informants for child personality: Agreement, discrepancies, and clinical utility. *Journal of Personality Assessment*, 93(6), 539–544. <http://dx.doi.org/10.1080/00223891.2011.608763>.
- Tackett, J. L., Herzhoff, K., Kushner, S. C., & Rule, N. (2015). Thin slices of child personality: Perceptual, situational, and behavioral contributions. *Journal of Personality and Social Psychology*. <http://dx.doi.org/10.1037/pspp0000044>.
- Van den Akker, A. L., Deković, M., Asscher, J., & Prinzie, P. (2014). Mean-level personality development across childhood and adolescence: A temporary defiance of the maturity principle and bidirectional associations with parenting. *Journal of Personality and Social Psychology*, 107(4), 736–750. <http://dx.doi.org/10.1037/a0037248>.
- Van der Giessen, D., Branje, S. J. T., Keijsers, L., Van Lier, P. C., Koot, H. M., & Meeus, W. H. J. (2014). Emotional variability during mother-adolescent conflict interactions: Longitudinal links to adolescent disclosure and maternal control. *Journal of Adolescence*, 37(1), 23–31. <http://dx.doi.org/10.1016/j.adolescence.2013.10.007>.
- Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, 98(2), 281–300. <http://dx.doi.org/10.1037/a0017908>.
- Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, 95(5), 1202–1216. <http://dx.doi.org/10.1037/a0013314>.
- Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology*, 78(3), 546–558. <http://dx.doi.org/10.1037/0022-3514.78.3.546>.

Watson, D., & Humrichouse, J. (2006). Personality development in emerging adulthood: Integrating evidence from self-ratings and spouse ratings. *Journal of Personality and Social Psychology*, *91*(5), 959–974. <http://dx.doi.org/10.1037/0022-3514.91.5.959>.

Weinert, F. E., & Schneider, W. (1999). *Individual development from 3 to 12: Findings from the Munich Longitudinal Study*. Cambridge, UK: Cambridge University Press.

Wood, A. M., Brown, G. D. A., Maltby, J., & Watkinson, P. (2012). How are personality judgments made? A cognitive model of reference group effects, personality scale responses, and behavioral reactions. *Journal of Personality*, *80*(5), 1275–1311. <http://dx.doi.org/10.1111/j.1467-6494.2012.00763.x>.