Description of meta-analytic procedure

Search procedure

We searched the Web of Knowledge database for articles with "personality" and "adolescence" in the title and obtained 268 hits. To be included in our analysis, the paper needed to report measures of the Big Five for at least two age groups between 10 and 20. This requirement was fulfilled by 9 papers with a total of 12 samples (Asendorpf & van Aken, 2003; Branje, van Lieshout, & Gerris, 2007; de Fruyt et al., 2006; de Leeuw, Scholte, Sargent, Vermulst, & Engels, 2010; Hill et al., 2013; Klimstra, Akse, Hale III, Raaijmakers, & Meeus, 2010; van den Akker, Deković, & Prinzie, 2010; van der Aa et al., 2009; Vecchione, Alessandri, Barbaranelli, & Caprara, 2012).

An additional 5 papers (containing 8 samples; Allik, Laidra, Realo, & Pullmann, 2004; De Haan, Deković, van den Akker, Stoltz, & Prinzie, in press; Klimstra, Crocetti, Hale, Fermani, & Meeus, 2011; Lounsbury, Hutchens, & Loveland, 2005; Poorthuis, Thomaes, Denissen, van Aken, & Orobio de Castro, 2012) were obtained from a supplementary Google Scholar search.

Coding of effect sizes

If a longitudinal study included multiple cohorts, these were counted as separate samples. Table 1 lists the resulting 20 samples. As can be seen, most came from publications that were only recently published, with a predominance of Belgian, Dutch, and Italian samples. Most samples were balanced with regard to gender and contained longitudinal data. Samples sizes ranged between 124 and 7888.

Effect sizes were computed based on published means and standard deviations. Specifically, the difference between two ages was divided by the SD of the first time point, thus computing a standardized difference (Cohen's *d*). In three cases (Allik et al., 2004; Branje et al., 2007; Poorthuis et al., 2012), the information was not contained in the article but were kindly provided by the first

authors. In one case (de Fruyt et al., 2006), the necessary information was not provided but an effect size could be computed using a formula to convert *F*-values to *d*-values.

If possible, separate effect sizes were calculated for boys and girls (see Table 1). A total of 60 effect sizes were thus obtained for each trait. These effect sizes indicated the amount of mean-level changes across an average retest interval (controlling for age) of 2.77 years. As displayed in Table 1, the included samples contributed between 1 and 12 effect sizes (i.e., when there were multiple age comparisons per longitudinal cohort, or multiple reported age groups per cross-sectional sample).

Results

Results were analyzed using the metafor (Version 1.7-0) meta-analysis package in R. This package does not allow for multiple dependent effects contributed by one source (e.g., a sample that is assessed at 3 or more time points). Therefore, we aggregated across samples, resulting in 20 effect sizes for each trait. These effect sizes were weighted by their variance (so that larger samples obtain more weight). Ages were also averaged across time lags, and dichotomized using the age 15 cut-off. Separate effect sizes were computed for both age groups. Whereas the effect sizes for the older age group did not statistically differ from zero, the average effect size for the younger age group was smaller than zero (i.e., indicating decreasing values across time) in the case of conscientiousness and openness, d = -0.07, Z = -2.03, p = .04 and d = -0.07, Z = -2.29, p = 0.02, respectively.

To further explore the shape of the age effects, we used multilevel modeling (using the nlme package, Version 3.1-108, in R), in which effect size (nested within sample) was predicted by age (recoded so that 0 = 10 years, 1 = 11 years, etc.) and age squared. Results indicated a significant linear age effect in the case of conscientiousness, $F(1,37) = 10.29$, p = .003. For openness, a linear as well as curvilinear age trend was found, $F(1,37) = 9.72$, p = .004, and $F(1,37) = 10.42$, p = .003, respectively.

In addition to replicating the results of the metafor procedure, two more subtle effects were found. Specifically, the multilevel analyses detected a statistically significant quadratic age effect for extraversion, $F(1,37) = 4.14$, $p < .05$, and a linear age effect for emotional stability, $F(1,37) = 4.31$, $p < .05$. Because these effects were not replicated in the meta-analysis, they were not further explored.

Graphical depiction

Results of the meta-analysis are depicted in the paper. These findings can be prone to misinterpretation, as they depict relative age differences (with the x-axis value as Time 1) instead of absolute values for a certain age. To facilitate interpretation, we saved the estimated values for both conscientiousness (using a linear time effect) and openness to experience (using a linear as well as a quadratic time effect). We then started from age 10 as a reference value and added the predicted slope value to this reference value. We then used the result of this computation and added the slope for age 12.77, age 15.54, and age 18.31 (i.e., multitudes of 2.77, the average time lag between two measurements within samples). For example, the predicted slope for openness for age 10 was -.33, so we plotted this as the predicted absolute value for age 12.77. For age 12.77, the predicted slope was -.04, so we plotted a value of -.33 + -.04 = -.37 for age 15.54.

The results are depicted in Figure 1 below and replicate the findings of a curvilinear mean-level development as also reported by Soto et al. (2011).

Figure 1.

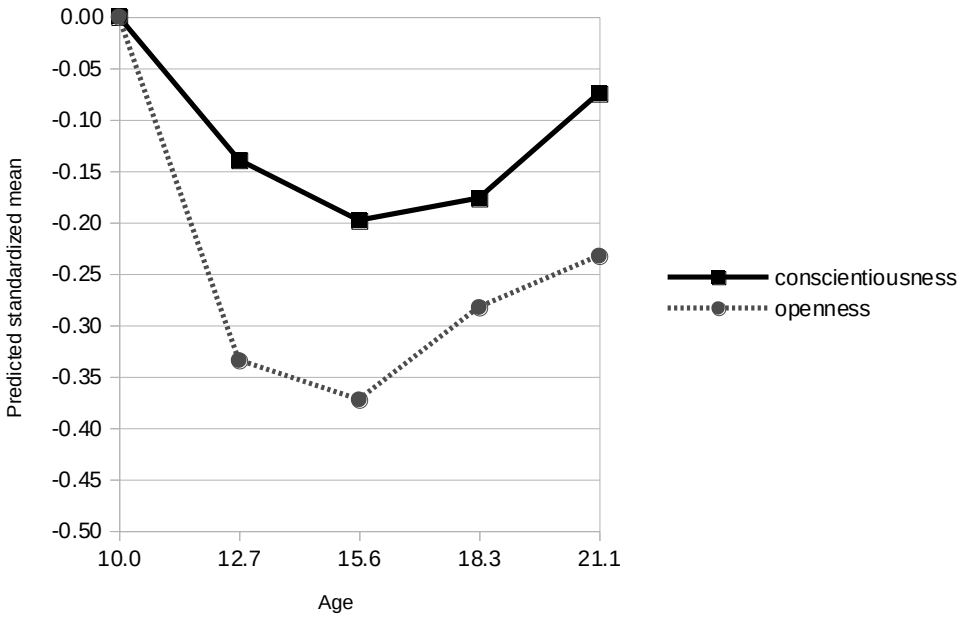
0.00
-0.05
-0.10
-0.15
-0.20
-0.25
-0.30
-0.35
-0.40
-0.45
-0.50
Predicted standardized mean
10.0
12.7
15.6
18.3
21.1
Age
conscientiousness
openness

Table 1.

Description of samples used in meta-analysis.

Sample	Publication year	Country	Sample (t1)	% Boys	Instrument	Longitudinal	results	Min age	Max age	Separate for gender	Data points (per trait)
Dehaan_inpress(a)	in press	BE	124	50	HiPIC	yes	Table 2	12.0	14.0	no	1
Dehaan_inpress(b)	in press	BE	139	50	HiPIC	yes	Table 2	10.0	15.0	no	2
Dehaan_inpress(c)	in press	BE	134	50	HiPIC	yes	Table 2	11.0	16.0	no	2
Dehaan_inpress(d)	in press	BE	132	50	HiPIC	yes	Table 2	12.0	17.0	no	2
Hill_2013	2013	CH	750	50	BFI-K	yes	Table 1	14.9	15.9	no	1
Poorthuis_2012	2012	NL	315	47	BFI	yes	Obtained	12.2	13.3	no	1
Vecchione_2012	2012	IT	403	48	Caprara	yes	Table 1	16.0	20.0	yes	4
Klimstra_2011b	2011	IT	1975	46	Gerris	no	Table 3	12.5	16.8	yes	2
Deleeuw_2010	2010	NL	428	50	Gerris	yes	Table 1	13.4	15.4	yes	4
Klimstra_2010(a)	2010	NL	923	47	Gerris	yes	Table 1	12.4	16.4	yes	8
Klimstra_2010(b)	2010	NL	390	47	Gerris	yes	Table 1	16.7	19.7	yes	8
Vandenakker_2010	2010	BE	290	49	HiPIC	yes	Table 1	11.8	14.8	no	1
Vanderaa_2009	2009	NL	7888	77	Gerris	no	Table 3	12.5	20.0	no	3
Branje_2006(a)	2006	NL	288	47	Gerris	yes	Obtained	12.4	14.4	no	2
Branje_2006(b)	2006	NL	288	50	Gerris	yes	Obtained	14.5	16.5	no	2
Defruyt_2006(a)	2006	BE	201	50	HiPIC	yes	Table 6	10.5	13.5	no	1
Defruyt_2006(b)	2006	BE	210	50	HiPIC	yes	Table 6	12.5	15.5	no	1
Lounsbury_2005	2005	US	851	49	APSI	no	Tables 1-3	12.6	17.8	no	2
Allik_2004	2004	EE	2650	46	NEO-FFI	no	Obtained	12.0	18.0	yes	12
Asendorpf_2003	2003	DE	230	52	Ostendorf	yes	Table 1	12.0	17.0	no	1