

Stability and Change in Adolescents' Personality: A Longitudinal Study

HELLE PULLMANN^{1,2*}, LIISA RAUDSEPP^{1,2} and JÜRI ALLIK^{1,2}

¹*Department of Psychology, University of Tartu, Tartu, Estonia*

²*The Estonian Centre of Behavioural and Health Sciences, Tartu, Estonia*

Abstract

The present study examined three types of personality change and continuity (mean-level, individual-level, and rank-order stability) over the 2-year period in a nationally representative longitudinal sample of Estonian adolescents (N = 876) aged 12–18. According to the Reliable Change Index, 82.1% of adolescents maintained the same level on any given personality trait measured by the NEO Five-Factorial Inventory (NEO-FFI) indicating that the individual-level continuity of adolescents did not differ compared to young adults. A reliable increase was found in Openness. Across the five dimensions, the average test–retest correlations were 0.51, 0.56 and 0.67, and the computed biennial stability values were 0.80, 0.83 and 0.89 for age groups 12 → 14, 14 → 16 and 16 → 18 years, respectively. Neither intelligence nor school performance moderated the differential continuity. Copyright © 2006 John Wiley & Sons, Ltd.

Key words: personality traits; NEO Five-Factorial Inventory (NEO-FFI); adolescence; temporal stability; longitudinal

INTRODUCTION

How stable are personality traits in the period of adolescence? Generally the patterns of changes in the personality traits of the five-factor model (FFM; McCrae & Costa, 1999) across adulthood are widely documented (e.g. Caspi, Roberts, & Shiner, 2005; Costa & McCrae, 2002; Roberts & DelVecchio, 2000). However, there are only few studies available that have examined change and continuity patterns in personality traits among adolescents using adults' self-report NEO questionnaires, extending thus previous knowledge of adult personality development based on the 'Big Five' questionnaires back to early adolescence (McCrae et al., 2002, as an example).

Several theoretical approaches to personality trait development have been proposed to date. At the broadest conceptual level, personality continuity and change may be attributed

*Correspondence to: Helle Pullmann, Department of Psychology, University of Tartu, Tiigi 78, 50410 Tartu, Estonia. E-mail: helle.pullmann@ut.ee

to nurture as biologically based (universal maturational processes) or to nurture as environmental influences. According to the five-factor model of personality, traits develop through childhood and reach maturity in adulthood showing little change for most personality traits after the age of 30 (McCrae & Costa, 2003). The observed changes in levels of personality traits are relatively slow and uniform across cultures suggesting minimal influence from external factors (McCrae et al., 2000, 1999). According to the contextual approach, mean-level changes in personality would result from one's transactions with the social environment (Zelli & Dodge, 1999), thus, being generally unpredictable. Recently, Roberts and Caspi (2003) proposed an alternative theory of personality trait development proposing that identity processes can help explain the patterns of continuity and change in personality traits across the life course.

The results of cross-sectional studies demonstrate that the mean levels of the 'Big Five' personality traits among adolescents resemble, quite closely, the respective scores of the adult population (Allik, Laidra, Realo, & Pullmann, 2004; Costa & McCrae, 2002). However, compared to adults, individual differences in traits are relatively fluid from age 12 to 18 years (McCrae et al., 2002). As Roberts and DelVecchio's (2000) review reported, the level of test-retest stability increases in a relatively linear fashion through adolescence and young adulthood. More (specifically, meta-analytic estimates of the mean population test-retest correlation coefficients showed that trait consistency increased from 0.31 in childhood to 0.45 in adolescence, to 0.54 during the college years, to 0.64 at age 30, and then reached a plateau around 0.74 between ages 50 and 70 when time interval was held constant at about 6–7 years (Roberts & DelVecchio, 2000). Thus, the magnitude of rank-order stability in childhood and adolescence, although not as high as among adults, is still remarkably high suggesting that a considerable part of individual differences remains relatively consistent during this time.

Although the general age trend for increasing stability is established, ambiguity remains about the change and continuity patterns in 'Big Five' personality traits during adolescence. This is caused by several factors including incommensurability of measuring instruments, accompanied by shortage of representative samples of adolescents. Relatively few studies have used a comprehensive set of personality variables to characterise young children and to track continuities and changes in their personalities over time, and many of them have serious limitations (Caspi et al., 2005). For example, in one of the most informative longitudinal studies a sample of gifted students was examined (McCrae et al., 2002), but the results have not been replicated in adolescent samples with a wider range of intellectual abilities. Although the description of one's own personality is not a very demanding task, it still requires some minimal amount of mental abilities and understanding of the relevant vocabulary (Allik, et al., 2004; Mõttus, Allik, & Pullmann, 2006). It is possible, for example, that the cognitive abilities and knowledge of young adolescents are not sufficient for sophisticated understanding of adults' personality questionnaire items. Therefore, researchers have developed specialised personality scales assuming that adult personality inventories are not necessarily appropriate for adolescents and children. As a result, data gathered by instruments that are used to study children's personalities are not directly comparable to adult findings from adult instruments (McCrae et al., 2002).

Are young adolescents still generally able to adequately evaluate and report their personality traits? There are several possibilities for disentangling personality dispositions from an individual's ability to evaluate them, for instance, (a) to include ratings of knowledgeable others or (b) to measure cognitive abilities that are required for observing one's own personality dispositions and for giving reliable self-reports on the basis of these observations. Previous studies have demonstrated that one's general reasoning level

contributes only to a certain level while evaluating one's own personality (Allik et al., 2004; Mõttus et al., 2006). For instance, Allik and his colleagues demonstrated that among 12-year-olds psychometrically measured intelligence predicted the reliability of self-reports and personality trait structure measured by the NEO Five-Factorial Inventory (NEO-FFI), indicating relevance to control for reasoning ability in studies of personality among young adolescents. However, recent studies have demonstrated that young adolescents have enough abilities to understand and respond properly to items from adults' self-report personality questionnaires (De Fruyt, Mervielde, Hoekstra, & Rolland, 2000; Markey, Markey, Tinsley, & Erickson, 2002; McCrae et al., 2002). Therefore, adult personality measures can be meaningfully used in adolescent samples while still bearing in mind a possible moderating effect of the intelligence level.

As personality traits are indisputably consistent across time and age, it is common to believe that if a construct demonstrates temporal consistency, it does not change (cf. Roberts, Walton, & Viechtbauer, 2006). In most longitudinal studies, consistency is operationalised as rank-order consistency, which refers to the relative placement of individuals within a group over time. Change is most often defined as mean-level change, which refers to whether a group of people increases or decreases on trait dimensions over time. The existence of consistency, at least as defined in terms of rank-order consistency, does not preclude the existence of change, especially mean-level change over time (Roberts et al., 2006). Therefore, it is important to study rank-order stability and mean-level change simultaneously in order to obtain a balanced overview of different aspects of personality stability and change during adolescence.

There are not many studies in which multiple cohorts, representative of the whole adolescent population, are followed longitudinally (cf. Robins, Fraley, Roberts, & Trzesniewski, 2001). Therefore, the longitudinal studies using representative samples with a wide age range would provide answers to the following basic questions concerning stability of personality traits before adulthood: How does the level of any given personality trait change over time when adolescents become older? Are there any individual differences in these changes? What is the differential continuity in adolescence and how does it depend on adolescents' age, gender, reasoning ability or academic achievement?

The goal of this study was to report three different aspects of personality continuity and change at a 2-year interval in a nationally representative longitudinal sample of adolescents aged 12–18. More specifically, the study focused on examining: (a) the mean-level changes in values of personality scores, (b) the individual-level continuity and (c) the rank-order stability in personality dimensions across three age groups (12 → 14, 14 → 16 and 16 → 18 years). In addition, psychometrically measured intelligence and academic achievement were regarded as possible moderators of differential continuity during adolescence.

METHOD

Participants

The data of this study were drawn from the Longitudinal Study of Estonian Schoolchildren. At Time 1 (2001), the sample of 2650 adolescents (1420 girls and 1230 boys) was tested (Allik et al., 2004, for details). Two years later (2003; Time 2), a nationally representative subsample of those adolescents ($N = 1383$) was invited to participate in the longitudinal study. As a result, a follow-up study was conducted in the chosen grades and complete answers for this

study were obtained from 876 school children (513 girls and 363 boys; mean age 16.1 years, $SD = 1.67$) attending Grades 8, 10 and 12. The retested sample was drawn from 17 Estonian-speaking secondary schools located in different regions of Estonia, covering all 15 counties, the capital city of Tallinn (11%), several towns (49%) and rural areas (40%). For further analyses, the retested adolescents were divided into three age groups (the youngest, middle and oldest) according to the mean ages in the tested grades at Time 2: (a) Age 12 → 14 ($n = 380$, $M = 14.5$ years, $SD = 0.56$), (b) Age 14 → 16 ($n = 206$, $M = 16.4$ years, $SD = 0.57$) and (c) Age 16 → 18 ($n = 290$, $M = 18.1$ years, $SD = 0.51$). The attrition rates between Time 1 and Time 2 were 22% and 27% for the youngest and oldest age groups, respectively. A high attrition rate (57%) in the middle age group (Age 14 → 16) was primarily caused by students' transition to another class or school (e.g. vocational schools) after completing the basic education level in Grade 9. Therefore, it was not possible to contact all those participants who had studied in the retested grades at Time 1. Testings were carried out during classes and all the students attending the grade at the testing day participated in the study. Consent to participate was obtained from adolescents and their parents.

Measures

Personality traits

Participants completed the Estonian version of the NEO-FFI; Allik et al., 2004; Costa & McCrae, 1992), which is a 60-item self-report measure of the five major personality domains: N, E, Openness to Experience (O), Agreeableness (A) and Conscientiousness (C). Each personality dimension is measured by 12 items on a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The internal reliability coefficients (Cronbach α) of the scales were 0.85 (N), 0.85 (E), 0.82 (O), 0.73 (A) and 0.84 (C) for the sample at Time 1. The mean internal reliability coefficients for age groups were 0.79 (Age 12 → 14), 0.81, (Age 14 → 16) and 0.84 (Age 16 → 18) at Time 2.

Intelligence

At Time 1 and Time 2, Raven Standard Progressive Matrices (SPM; Raven, 1981) test was administered without any time limits to measure the participants' non-verbal reasoning ability. The internal reliabilities of the SPM ranged from $\alpha = 0.82$ to 0.88 for the age groups. The SPM was standardised in Estonia on the total sample of participants at Time 1 (Lynn, Allik, Pullmann, & Laidra, 2002).

Academic achievement

At Time 1 and Time 2, academic achievement was measured using the students' grade point average (GPA) in academic subjects reported on the school record from the last semester or two previous quarters. A 5-point grading system ranging from 1 (*very poor*) to 5 (*very good*) is currently applied in Estonia.

RESULTS

Mean-level stability

The mean-level stability examines the degree to which the participants as a group have changed in their self-reported personality traits from Time 1 to Time 2. In Table 1, the

Table 1. Mean-level stability and reliable change in the NEO personality traits across the age groups over the 2-year period

Age	Scale	Time 1		Time 2		<i>d</i>	Individual-level change		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		Decreased (%)	No reliable change (%)	Increased (%)
12 → 14 [†]	N	49.8	9.17	48.5	8.98	-0.15**	12.4	78.8	8.9
	E	52.3	8.07	54.3	8.65	0.24***	6.1	79.7	14.2
	O	41.4	7.63	42.2	8.23	0.10	6.6	85.2	8.2
	A	53.4	9.91	51.9	10.11	-0.15**	11.6	82.6	5.8
	C	51.7	8.11	52.2	8.08	0.06	8.4	82.4	9.2
14 → 16 [‡]	N	50.8	9.21	49.6	10.19	-0.13*	14.1	80.1	5.8
	E	53.3	9.36	54.6	9.31	0.14*	6.8	80.6	12.6
	O	41.4	8.94	43.6	8.65	0.25***	8.3	74.7	17.0
	A	50.0	10.33	50.3	10.61	0.03	8.3	81.0	10.7
	C	51.1	8.84	50.8	9.80	-0.03	8.4	81.4	10.2
16 → 18 [#]	N	49.2	8.87	48.9	9.79	-0.04	9.3	85.9	4.8
	E	53.3	9.67	53.1	10.74	-0.02	9.7	82.8	7.5
	O	43.0	8.41	45.4	9.69	0.27***	3.4	82.8	13.8
	A	48.8	8.82	49.2	10.41	0.04	6.6	87.5	5.9
	C	50.0	8.85	50.7	9.20	0.08	5.5	86.2	8.3
Total	N	49.8	9.09	48.9	9.55	-0.11***	11.9	81.6	6.5
	E	52.9	8.94	53.9	9.55	0.12***	7.5	81.0	11.4
	O	41.9	8.24	43.6	8.94	0.19***	6.1	80.9	13.0
	A	51.1	9.87	50.7	10.39	-0.04	8.8	83.7	7.5
	C	51.0	8.56	51.4	8.90	0.04	7.4	83.3	9.2

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

[†] $n = 380$; [‡] $n = 206$; [#] $n = 290$.

T1 = Time 1 (2001), *T2* = Time 2 (2003). *M* = mean, *SD* = standard deviation, *d* = mean-level change in standard score units (Cohen's *d*); N = Neuroticism, E = Extraversion, O = Openness to Experience, A = Agreeableness, C = Conscientiousness.

difference between mean levels of both testing times is reported using the standardised mean differences (Cohen's *d*); significance of the change was based on the dependent samples *t*-tests. The mean scores of the NEO-FFI are expressed as *T* scores ($M = 50$, $SD = 10$) based on within-sex norms of the young Estonian adults (see Allik et al., 2004, for details).

Figure 1 demonstrates the mean levels of the NEO-FFI domain scores for the age groups at two times. The results indicated that among the youngest age group (Age 12 → 14) the level of Extraversion has increased on average from 52.3 to 54.3 ($d = 0.24$, $p < 0.001$) during the 2 years. Also significant mean level decreases were found on the Neuroticism and Agreeableness scales for these young adolescents. Although the Openness dimension did not reveal any significant mean-level differences in the youngest group, on average, Openness level changed the most across all participants ($d = 0.19$, $p < 0.001$), showing the largest increase for the oldest and middle age groups ($d = 0.25$ and 0.27 , respectively, $p < 0.001$).

Gender differences were found in the Extraversion dimension in which boys did not show any significant differences in the mean levels of their E score compared to girls in the

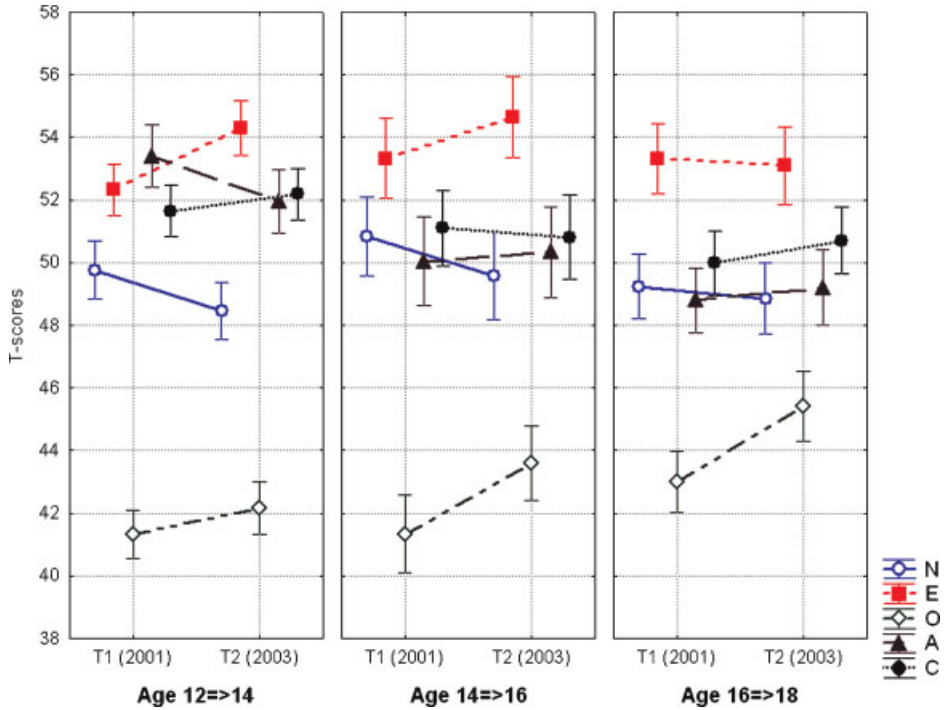


Figure 1. The mean levels of the NEO-FFI domain scores for the age groups at Time 1 (T1) and Time 2 (T2). Vertical bars denote 0.95 confidence intervals. N=Neuroticism, E=Extraversion, O=Openness, A=Agreeableness, C=Conscientiousness.

youngest age group; in the oldest age group the differences were statistically significant. In addition, girls' level of Neuroticism remained the same in the youngest group while boys became emotionally more stable over the 2-year period.

Individual-level change

The Reliable Change Index (RCI; Christensen & Mendoza, 1986; Jacobson & Truax, 1991) was calculated to reveal how many participants have changed in any given trait more than would be expected considering the unreliability of the NEO-FFI. For this purpose, the observed difference scores were compared with the distribution of change scores that would be expected from error of measurement alone. Scores that exceeded a 95% confidence interval were assumed to represent true increase or decrease on the scale and, based on these scores, all participants were divided into three groups ('increased', 'decreased' and 'no reliable change') broken down by the personality traits.

The majority of adolescents did not demonstrate a reliable change over the 2-year period as, on average, 82.%1 (ranging from 74.7 to 87.5) of participants maintained the same level in their personality traits across time (Table 1). The most remarkable changes were found in the youngest group (Age 12 → 14) with the maximum increase of 14.2% in Extraversion and decrease of 12.4% in Neuroticism. Generally, the level of Openness had changed the most with an increase of 13.0% while about 11.9% of the adolescents had a reliably reduced level of Neuroticism according to the RCI.

Rank-order stability

The *NEO-FFI* test-retest correlation coefficients over the 2-year period in the three age groups of adolescents are reported separately for girls and boys in Table 2. The results indicated that the degree to which the relative differences among individuals remained stable over time increased as a function of age. The average test-retest correlations across all five personality dimensions were 0.51, 0.56 and 0.67 for the age groups 12 → 14, 14 → 16 and 16 → 18, respectively.

Girls were generally slightly more stable in their ordering on personality dimensions compared to boys and their test-retest correlations increased gradually as a function of age. This pattern was similar for boys on the E and C dimensions but interestingly, the rank-order stability of N remained at the same level across the age groups. Moreover, male students dropped notably in the stability indexes of the O and A scales during their transition from Grade 8 to 10; the difference in the test-retest correlations in Agreeableness was statistically significant ($p < 0.05$).

The rank order stability can be also partitioned into individual contributions to overall rank-order consistency. The individual stability coefficient can be defined as

$$1 - \frac{(z_1 - z_2)^2}{2},$$

where z_1 and z_2 are scores for a trait standardised across the sample at the first and second administrations (Asendorpf, 1992). The mean of Asendorpf's coefficient across all participants is equal to the test-retest correlation. The individual stability coefficients were unsystematically related to the mean level of personality traits reported either during the first or second administration. Only stability on Extraversion was related to the mean level of Neuroticism and Extraversion: more consistent children were emotionally more stable [$r(870) = -0.09$ on the both administrations, $p < 0.01$] and extraverted [$r(857) = 0.12$ and 0.14 on the first and second administration, respectively; both $p < 0.001$]. When individual stability coefficients on all five personality dimensions were correlated with individual increase in the SPM scores only consistency in Extraversion demonstrated a modest association in the total sample: those adolescents who were more consistent in reporting

Table 2. Rank-order stability in the NEO personality traits across the age groups and sexes over the 2-year period

	Age $T1 \rightarrow T2$	n	N	E	O	A	C	M
Girls	12 → 14	215	0.49	0.50	0.45	0.49	0.46	0.48
	14 → 16	120	0.68	0.64	0.47	0.61	0.67	0.61
	16 → 18	178	0.72	0.73	0.67	0.61	0.70	0.69
	Total	513	0.61	0.63	0.55	0.57	0.60	0.59
Boys	12 → 14	165	0.52	0.39	0.53	0.48	0.46	0.48
	14 → 16	86	0.52	0.52	0.38	0.34	0.48	0.45
	16 → 18	112	0.49	0.71	0.60	0.60	0.66	0.61
	Total	363	0.51	0.55	0.52	0.48	0.53	0.52
Total	12 → 14	380	0.50	0.49	0.57	0.51	0.48	0.51
	14 → 16	206	0.64	0.62	0.48	0.49	0.59	0.56
	16 → 18	290	0.65	0.73	0.68	0.62	0.69	0.67
	Total	876	0.58	0.61	0.59	0.54	0.59	0.58

n = Number of participants; N = Neuroticism, E = Extraversion, O = Openness to Experience, A = Agreeableness, C = Conscientiousness, M = the mean test-retest correlation. All correlations are significant at $p < 0.001$.

their level of Extraversion on two successive administrations had a slightly larger increase in their mental development [$r(862) = 0.07, p = 0.04$]. However, this result was not found separately for the age groups. There were no significant correlations between academic achievement and individual stability coefficients on all five personality dimensions.

Moderation of personality stability

Intelligence and academic achievement were considered as two potential moderators of stability of personality traits at Time 1 and Time 2. To examine potential moderation, the SPM and GPA scores were partialled out from personality test–retest correlations. The NEO-FFI test–retest correlations were computed when (a) the SPM scores or (b) the GPA were taken into account at Time 1 or at Time 2. Controlling for the intelligence test score had a relatively small effect upon the test–retest correlations across all the personality dimensions. Any of the test–retest correlations for the age groups did not change statistically significantly when any of the moderators (SPM and GPA, both at Time 1 and Time 2) were taken into account and changes in correlations did not exceed 0.05 in magnitude.

Another possibility to test the moderating effects is to compare stability coefficients in high and low groups, using a median split on the moderator variable. Therefore, all participants were divided within each age group into two subgroups on the basis of a median split of the SPM score. The results did not reveal any statistically significant ($p < 0.05$) differences between the test–retest correlations in low and high intelligence groups for this sample. Thus, the results confirmed that adolescents with lower level of reasoning ability did not have more inconsistencies in their self-descriptions compared to students with higher general mental abilities. The same negative result was obtained when a multiplicative term (personality scores at Time 1 \times SPM/GPA scores) was added to the multiple regression: personality scores at Time 2 were not better predicted when multiplicative terms were added.

DISCUSSION

As a whole, the current study has made several contributions to the literature on personality development. Firstly, a nationally representative longitudinal sample of adolescents aged 12–18 was used. Several studies have been carried out before, examining mean-level changes in the ‘Big Five’ personality traits in adolescence cross-sectionally (e.g. De Fruyt et al., 2000). However, the available longitudinal papers in personality development are commonly not based on representative adolescents’ samples (e.g. McCrae et al., 2002; Parker & Stumpf, 1998). The participants in the present longitudinal study represented a wide range of intellectual abilities and various socioeconomic backgrounds covering the age range from 12 to 18 years. Secondly, among adult samples a well-validated ‘Big Five’ questionnaire was used to assess students’ self-reported personality dimensions. The data gathered by the NEO-FFI among adolescents in this study are directly comparable to adult findings based on the NEO instruments, extending thus the previous knowledge of adult personality development to early adolescence. Thirdly, three different aspects of personality continuity and change were reported across a 2-year interval. There are only few studies available which have examined several continuity types simultaneously (e.g. De Fruyt, Van Leeuwen, Bagby, Rolland, & Rouillon, 2006; Roberts, Caspi, &

Moffitt, 2001; Robins et al., 2001, as examples). Fourthly, cross-cultural contribution of the present study should also be emphasised because of a non-English-speaking sample. Below, we will discuss briefly some main results of the current study.

Mean-level stability

In our previous study of the same sample of adolescents we demonstrated that the mean-levels of their personality traits were quite similar to the respective scores of Estonian adults (Allik et al., 2004). Moreover, all traits demonstrated only a modest cross-sectional change in the mean-level of the trait scores: the level of Openness increased and the levels of Agreeableness and Conscientiousness decreased between 12 and 18 years of age (Allik et al., 2004). The results of a recent study of another longitudinal sample of Estonian adolescents confirmed that the mean level of 'Big Five' personality traits remained more or less on the same level from 15 to 18 years of age also when reported by mothers, fathers and teachers (Laidra, Allik, Harro, Merenäkk, & Harro, 2006). However, analyses of the current longitudinal study confirmed that the increase of the mean level of Openness was observable both in the cross-sectional and longitudinal data. Therefore we can conclude that from the age of 14 years onwards adolescents become more tolerant and open to new ideas and experiences. In addition, the longitudinal data indicated an increasing tendency to become more emotionally stable. Follow-up data, however, did not confirm the previous cross-sectional finding (Allik et al., 2004) that adolescents as a group become on average less agreeable and less conscientious over years. To conclude, the increase in Openness was a reliable developmental change, although being rather modest by its magnitude.

Individual-level change

Mean-level change occurs when most adolescents change in the same way over time. Thus, it is possible, for example, that some adolescents became more extraverted and socially agreeable but approximately the same number of adolescents became more introverted and antagonistic. In this case, changes in mean-levels of the traits are minimal as opposite tendencies are offsetting each other. However, there may still be individual differences in the change reflected by changes in the individual scores. The results of this study demonstrated that the absolute majority of individual scores in about 82% cases remained on the same level concerning any of the five dimensions of personality over the 2-year period in this sample. Moreover, this percentage was practically the same in each age group during adolescence. Therefore, young people aged 12–18 report their personality traits generally in a consistent manner using self-report inventories. This result is in line with a recent study (Laidra, Allik, Harro, Merenäkk, & Harro, 2006) in which only 10–15% of parents reported that the mean level of the 'Big Five' personality traits of their teenage child had changed significantly over the 3-year period.

Do individual scores change more during the period of adolescence than among young adults? Robins and his colleagues (2001) found that among young adults 82.4% of the participants remained the same level on any given trait measured by the NEO-FFI over a 4-year period. Vaidya, Gray, Haig, and Watson (2002) studied young adults over a 2.5-year period and found that in 83.2% of the participants stayed the same level on the Big Five Inventory (BFI). Roberts, Caspi, and Moffit (2001) analysed individual-level personality continuity from 18 to 26 years in a birth cohort ($N = 921$) using the Multidimensional Personality Questionnaire. They reported that 72–84% of their participants, depending on

the measured attribute, remained stable over the 8-year interval between the two administrations. To conclude, the results of the current study demonstrate that individual-level stability of traits among adolescents does not differ remarkably compared to young adults.

Rank-order stability

Continuity and change are most often indexed by correlations between personality scores across two points in time (Caspi et al., 2005). These differential or rank-order stability correlations reflect the degree to which the relative ordering of individuals on a given trait is maintained over time.

According to a meta-analytic survey by Roberts and DelVecchio (2000), a typical mean value of test-retest for the 12- to 18-year-olds is about 0.45 when the time interval was held constant at 6.7 years. It has been demonstrated that the observed test-retest correlation (C) can be expressed as a product of the internal consistency (R) and the annual stability (s) rose to a power of years (n) over which the test-retest stability was calculated (Conley, 1984; Converse & Markus, 1979):

$$C = Rs^n$$

Knowing that the average internal consistency was 0.74 (Roberts & DelVecchio, 2000), the estimated annual stability (s) equals 0.93, which remained below the typical value of the annual stability for adults. For example, Conley (1984) found that the annual stability of adult personality traits might be estimated as high as 0.98. However, the annual stability is not invariant of the measuring instrument and the shorter scales have typically lower annual stability.

McCrae and his colleagues (2002) reported a 4-year longitudinal study of gifted American students ($N = 230$) who completed the NEO-FFI for the first time when they were, on average, 12 years old and again at the age of 16. The 4-year test-retest correlation was 0.40 and 0.41 for boys and girls, respectively. Knowing that the average internal consistency (Cronbach alpha) was 0.78 from the first testing (Parker & Stumpf, 1998), we can estimate the annual stability at 0.85. As the same instrument was applied in this study, we can use this estimate as a reference point. In this study, the mean test-retest values across the five personality dimensions were 0.51, 0.56 and 0.67 (Table 2) for age groups 12 → 14, 14 → 16 and 16 → 18 years, respectively. Based on the mean internal stability coefficients of 0.79, 0.81 and 0.84, the biennial stability values of 0.80, 0.83 and 0.89 can be computed for these age groups, respectively. Thus, the biennial stability in self-reported personality descriptions of 16-year old Estonians (0.83) resembles quite closely the annual stability of the gifted American children of the same age (0.85) reported by McCrae and his colleagues (2002).

As another comparison, Vaidya and his colleagues (2002) used the BFI to study personality stability among university students with an approximately 2.5-year interval between two testings. The BFI has 8–10 items measuring each dimension, which is comparable to the length of the NEO-FFI scales (each dimension is represented by 12 items). The mean reliability across all dimensions was 0.81 and the mean test-retest correlation 0.64 (Vaidya et al., 2002). From these two values we can estimate the annual stability, which is 0.91 for students who were on average 21 year old during the second testing. This value is only slightly higher than the annual stability observed in the oldest age

group of the present study suggesting that 16-year-olds have already reached the rank-order stability that is comparable to young adults. Although the test-retest correlations were smaller for younger adolescents than those that are typical for young adults, the values were still remarkably high.

In addition, the possible moderating effect of general intelligence on the differential continuity was explored as the period of adolescence can be described by rapid development of intellectual abilities (Pullmann, Allik, & Lynn, 2004). The need to control for intelligence level in this study was based on a recent study (Allik et al., 2004), which demonstrated that among young adolescents non-verbal intelligence predicted the reliability of self-reports and personality trait structure measured by the NEO-FFI. In this study, it was hypothesised that young adolescents with a higher level of mental abilities might understand personality items in a more comprehensive way and give more reliable self-reports compared to students with a lower level of intelligence, which could have an effect on consistency and stability of their self-report descriptions. Although general intelligence is significantly related to academic achievement (Laidra, Pullmann, & Allik, 2006, for this sample), performance in school context also depends on various other aspects like social relations, environment etc. As there is a significant link between personality traits and achievement in school context (Caspi et al., 2005, for an overview) a possible mediating effect of academic achievement in self-reported personality traits among adolescents was also tested.

Based on the results of this study, we can conclude that neither intelligence nor academic achievement moderate stability of personality traits in adolescence. In other words, adolescents' reasoning ability or success in school does not have an effect upon how stable they are in their self-descriptions over the 2-year period. As a limitation, it was not possible to include other moderators in the analyses as there was no information available about adolescents' health, family relationships (Branje, van Lieshout, & van Aken, 2004), life experiences (Vaidya et al., 2002), private self-consciousness (McCrae, 1993) or other variables that have been studied as moderators of differential stability of personality in previous longitudinal studies. Nevertheless, the present study extended the knowledge of the continuity and change of the five-factor model of personality traits to the whole period of adolescence from ages 12 to 18.

ACKNOWLEDGEMENTS

This project was supported by the Estonian Science Foundation (Grants 4519 and 5677) and the Estonian Ministry of Science and Education (Grant 0180543). Parts of this article were presented at the 12th European Conference on Personality, July 2004, Groningen, The Netherlands. The authors thank Richard Lynn, Kaia Laidra and Anneli Veisson for their help in carrying out this project. The authors are also grateful to Jeff McCrae and two reviewers for their helpful comments on this manuscript.

REFERENCES

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

- Asendorpf, J. B. (1992). Beyond stability: Predicting inter-individual differences in intra-individual change. *European Journal of Personality*, 6, 103–117.
- Branje, S. J. T., van Lieshout, C. F. M., & van Aken, M. A. G. (2004). Relations between Big Five personality characteristics and perceived support in adolescents' families. *Journal of Personality and Social Psychology*, 86, 615–628.
- Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, 56, 453–484.
- Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behavior Therapy*, 17, 305–308.
- Conley, J. J. (1984). The hierarchy of consistency: A review and model of longitudinal findings on adult individual differences in intelligence, personality and self-opinion. *Personality and Individual Differences*, 5, 11–25.
- Converse, P. E., & Markus, G. B. (1979). Plus ça change . . . : The new CPS election study panel. *American Political Science Review*, 73, 32–49.
- Costa, P. T., Jr., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI). In *Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., Jr., & McCrae, R. R. (2002). Looking backward: Changes in the mean levels of personality traits from 80 to 12. In D. Cervone, & W. Mischel (Eds.), *Advances in personality science* (pp. 219–237). New York: Guilford.
- De Fruyt, F., Mervielde, I., Hoekstra, H. A., & Rolland, J. P. (2000). Assessing adolescents' personality with the NEO PI-R. *Assessment*, 7, 329–345.
- De Fruyt, F., Van Leeuwen, K., Bagby, R. M., Rolland, J. P., & Rouillon, F. (2006). Assessing and interpreting personality change and continuity in patients treated for major depression. *Psychological Assessment*, 18, 71–80.
- Jacobson, J. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19.
- Laidra, K., Allik, J., Harro, M., Merenäk, L., & Harro, J. (2006). Agreement among adolescents, parents, and teachers on adolescent personality. *Assessment*, 13, 187–196.
- Laidra, K., Pullmann, H., & Allik, J. (2006). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences*, DOI:10.1016/j.paid.2006.08.001
- Lynn, R., Allik, J., Pullmann, H., & Laidra, K. (2002). A study of intelligence in Estonia. *Psychological Reports*, 91, 1022–1026.
- Markey, P. M., Markey, C. N., Tinsley, B. J., & Ericksen, A. J. (2002). A preliminary validation of preadolescents' self-reports using the Five-Factor Model of personality. *Journal of Research in Personality*, 36, 173–181.
- McCrae, R. R. (1993). Moderated analyses of longitudinal personality stability. *Journal of Personality and Social Psychology*, 65, 577–585.
- McCrae, R. R., & Costa, P. T. (1999). A five-factor theory of personality. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality theory and research* (Vol. 2, pp. 139–153). New York: Guilford Press.
- McCrae, R. R., & Costa, P. T. (2003). *Personality in adulthood* (2nd ed.). New York: Guilford Press.
- McCrae, R. R., Costa, P. T., Ostendorf, F., Angleitner, A., Hrebickova, M., Avia, M. D., Sanz, J., Sanchez-Bernardos, M. L., Kusdil, M. E., Woodfield, R., Saunders, P. R., & Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology*, 78, 173–186.
- McCrae, R. R., Costa, P. T., Jr., Pedroso de Lima, M., Simoes, A., Ostendorf, F., Angleitner, A., Marusic, I., Bratko, D., Caprara, G. V., Barbaranelli, C., Chae, J. H., & Piedmont, R. L. (1999). Age differences in personality across the adult life span: Parallels in five cultures. *Developmental Psychology*, 35, 466–477.
- McCrae, R. R., Costa, P. T., Jr., Terracciano, A., Parker, W. D., Mills, C. J., De Fruyt, F., & Mervielde, I. (2002). Personality trait development from age 12 to age 18: Longitudinal, cross-sectional, and cross-cultural analyses. *Journal of Personality and Social Psychology*, 83, 1456–1468.

- Mõttus, R., Allik, J., & Pullmann, H. (2006). Does personality vary across ability levels: A study using self and other ratings. *Journal of Research in Personality*. DOI:10.1016/j.jrp.2006.03.007
- Parker, W. D., & Stumpf, H. (1998). A validation of the Five-Factor Model of personality in academically talented youth across observers and instruments. *Personality and Individual Differences*, 25, 1005–1025.
- Pullmann, H., Allik, J., & Lynn, R. (2004). The growth of IQ among Estonian schoolchildren from 7 to 19. *Journal of Biosocial Science*, 36, 735–740.
- Raven, J. (1981). *Manual for Raven's Progressive Matrices and Mill Hill Vocabulary Scales*. Oxford: Oxford Psychologists.
- Roberts, B. W., & Caspi, A. (2003). The cumulative continuity model of personality development: Striking a balance between continuity and change in personality traits across the life course. In U. Staudinger, & U. Lindenberger (Eds.), *Understanding human development: Lifespan psychology in exchange with other disciplines* (pp. 183–214). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Roberts, B. W., Caspi, A., & Moffitt, T. E. (2001). The kids are alright: Growth and stability in personality development from adolescence to adulthood. *Journal of Personality and Social Psychology*, 81, 670–683.
- Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, 126, 3–25.
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, 132, 1–25.
- Robins, R. W., Fraley, R. C., Roberts, B. W., & Trzesniewski, K. H. (2001). A longitudinal study of personality change in young adulthood. *Journal of Personality*, 69, 617–640.
- Vaidya, J. G., Gray, E. K., Haig, J., & Watson, D. (2002). On the temporal stability of personality: Evidence for differential stability and the role of life experiences. *Journal of Personality and Social Psychology*, 83, 1469–1484.
- Zelli, A., & Dodge, K. A. (1999). Personality development from the bottom up. In D. Cervone, & Y. Shoda (Eds.), *The coherence of personality: Social-cognitive bases of consistency, variability, and organization* (pp. 94–126). New York: Guilford Press.