

The Little Six Personality Dimensions From Early Childhood to Early Adulthood: Mean-Level Age and Gender Differences in Parents' Reports

Christopher J. Soto

Colby College

Abstract

The present research pursues three major goals. First, we develop scales to measure the Little Six youth personality dimensions: Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Activity. Second, we examine mean-level age and gender differences in the Little Six from early childhood into early adulthood. Third, we examine the development of more specific nuance traits. We analyze parent reports, made using the common-language California Child Q-Set (CCQ), for a cross-sectional sample of 16,000 target children ranging from 3 to 20 years old. We construct CCQ–Little Six scales that reliably measure each Little Six dimension. Using these scales, we find (a) curvilinear, U-shaped age trends for Agreeableness, Conscientiousness, and Openness, with declines followed by subsequent inclines; (b) monotonic, negative age trends for Extraversion and Activity; (c) higher levels of Conscientiousness and Agreeableness among girls than boys, as well as higher levels of Activity among boys than girls; and (d) gender-specific age trends for Neuroticism, with girls scoring higher than boys by mid-adolescence. Finally, we find that several nuance traits show distinctive developmental trends that differ from their superordinate Little Six dimension. These results highlight childhood and adolescence as key periods of personality development.

Several key patterns have emerged in the study of life span personality development: Children, adolescents, and adults can all be described in terms of personality traits—characteristic patterns of thinking, feeling, and behaving (Caspi, Roberts, & Shiner, 2005). Personality traits do not become fixed at any particular age; they remain capable of change throughout the life span (Roberts & DelVecchio, 2000). Most adults become more agreeable, conscientious, and emotionally stable as they age, a phenomenon dubbed the maturity principle (Roberts, Walton, & Viechtbauer, 2006; Roberts & Wood, 2006). There are modest mean-level gender differences in personality: In general, women tend to be somewhat more extraverted, agreeable, conscientious, and neurotic than men, although these differences vary across cultures (e.g., De Bolle et al., 2015; Schmitt, Realo, Voracek, & Allik, 2008).

These points of consensus constitute major advances in our understanding of personality development, but they also raise new questions. For example, much more is known about normative personality development in adulthood than childhood (Caspi et al., 2005). How do youths' personality traits typi-

cally develop? Does the adult trend toward greater psychosocial maturity extend backward into childhood and adolescence? When and how do gender differences in personality first emerge? The present research addressed these questions by examining mean-level age and gender differences in parent-reported personality traits across early childhood (which we define as approximately ages 3–5), middle childhood (ages 6–9), late childhood (ages 10–12), early adolescence (ages 13–14), late adolescence (ages 15–17), and into early adulthood (ages 18–20).

The author thanks Grace DiBri and Ivan Yang for their assistance with this research, and Oliver P. John for providing the complete principal component loading matrix from John, Caspi, Robins, Moffitt, and Stouthamer-Loeber (1994).

Correspondence concerning this article should be addressed to Christopher J. Soto, Colby College, Department of Psychology, 5550 Mayflower Hill, Waterville, ME 04901. Email: christopher.soto@colby.edu.

Temperament, Personality, and the Little Six

Progress toward understanding personality development in childhood and adolescence has been slowed by the distinction historically drawn between child temperament and adult personality (Caspi et al., 2005). Temperament is often defined as behavioral and affective traits that appear within the first few years of life and have a strong biological basis (Goldsmith et al., 1987). In contrast, personality has been thought to gradually emerge over the course of childhood and adolescence, as temperamental dispositions become psychologically elaborated into personality traits (Rothbart, 2007; Shiner & Caspi, 2003). Reflecting this distinction, temperament and personality are typically measured using different instruments that assess different sets of traits. Temperament models and measures most often include versions of four trait dimensions: surgency/sociability (vs. shyness/inhibition), negative emotionality, persistence/effortful control (vs. impulsivity), and activity level (Buss & Plomin, 1984; De Pauw & Mervielde, 2010; Rothbart, Ahadi, Hershey, & Fisher, 2001; Thomas & Chess, 1977). In contrast, personality is most commonly assessed in terms of the Big Five trait dimensions: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience (Goldberg, 1990; John, Naumann, & Soto, 2008; McCrae & Costa, 1987).

A number of recent reviews have worked toward connecting the temperament and personality literatures by highlighting conceptual and empirical overlaps between them (e.g., Caspi et al., 2005; Clark & Watson, 2008; Shiner & DeYoung, 2013). For example, the temperament traits of surgency/sociability, negative emotionality, and persistence/effortful control have clear parallels with the personality traits of Extraversion, Neuroticism, and Conscientiousness, respectively. Drawing on these links, Shiner and DeYoung (2013) and Soto and John (2014) recently proposed that basic individual differences in youths' psychological characteristics may be best conceptualized in terms of six broad trait dimensions: Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Activity.

This "Little Six" model represents a conceptual union of the most prominent traits in the temperament and personality literatures. Preliminary empirical support for the model comes from a study examining the joint structure of several temperament and personality measures (De Pauw, Mervielde, & Van Leeuwen, 2009), as well as research examining the multidimensional structure of the California Child Q-Set (CCQ; Block & Block, 1980), a broadband measure of youths' personal characteristics (John, Caspi, Robins, Moffitt, & Stouthamer-Loeber, 1994; Soto & John, 2014; Van Lieshout & Haselager, 1994). Each of these studies identified a multidimensional structure that included all of the Little Six as independent dimensions.

The Little Six model thus holds promise for both describing youths' traits and integrating the temperament and personality literatures. However, further investigation of this model is impeded by the lack of measures that independently assess each Little Six dimension. Therefore, the present research's first

major goal was to develop a method for assessing the Little Six using the item pool of the common-language CCQ (Block & Block, 1980; Caspi et al., 1992). Previous research has shown that the CCQ can be used to measure youth versions of Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness (John et al., 1994), and that Activity is largely independent of these dimensions in childhood (e.g., De Pauw et al., 2009; Soto & John, 2014; Van Lieshout & Haselager, 1994). We therefore expected that the CCQ would include enough content to reliably measure each Little Six personality dimension.¹

Age and Gender Differences in the Little Six

Important biological, social, and psychological changes occur throughout childhood and adolescence. Biologically, youths change in terms of body size and shape, hormone levels, and brain anatomy and chemistry (Keating, 2004; Marshall & Tanner, 1986). Socially, their relationships with parents and peers evolve; romantic relationships emerge and become increasingly important (Collins, 2003; Rice & Mulkeen, 1995). Psychologically, they gain new cognitive, emotional, and behavioral capacities, and they work to develop coherent and differentiated identities (Erikson, 1968; Harter, 2006; Inhelder & Piaget, 1958).

What pattern of personality development might these changes produce? One possibility, which we will refer to as the maturity hypothesis, is that the positive age trends in personality development often observed during adulthood might extend backward into childhood and adolescence. That is, youths may become steadily more agreeable, more conscientious, and more emotionally stable across childhood and adolescence. However, other hypotheses are also plausible. One alternative, which we term the *disruption hypothesis*, proposes that some of the biological, social, and psychological changes experienced during the transition from childhood to adolescence may produce adjustment problems and might therefore be accompanied by temporary dips in psychosocial maturity.

One large cross-sectional study of youths' self-reports provides some support for the disruption hypothesis (Soto, John, Gosling, & Potter, 2011). Specifically, this study found curvilinear, U-shaped age trends for Agreeableness, Conscientiousness, and Openness. These traits declined from late childhood into early adolescence, and then inclined from late adolescence into adulthood.² This study also found declines in Extraversion and Activity (measured using a subset of Extraversion items) from late childhood into adolescence, followed by flat age trends through adulthood. Additional support for the disruption hypothesis comes from a recent longitudinal study spanning from middle childhood into early adulthood (Van den Akker, Deković, Asscher, & Prinzie, 2014), as well as a meta-analysis of 14 cross-sectional and longitudinal studies that collectively spanned late childhood and adolescence (Denissen, Van Aken, Penke, & Wood, 2013). Like Soto et al. (2011), both studies found U-shaped age trends for Conscientiousness and

Openness, as well as a decline in Extraversion; Van den Akker et al. (2014) also found a U-shaped trend for Agreeableness.

There are also reasons to suspect that the development of some personality traits might differ by gender. In adulthood, mean levels of Extraversion, Agreeableness, Conscientiousness, and Neuroticism tend to be somewhat higher among women than men (De Bolle et al., 2015; Schmitt et al., 2008), but it is not yet clear when these gender differences first emerge. Childhood and adolescence tend to be experienced differently by boys versus girls (Dweck, 1986; Hill & Lynch, 1983), and these differences may influence personality development. For example, Soto et al. (2011) found that, from late childhood into adolescence, mean levels of Neuroticism inclined among girls but not boys, producing a substantial gender difference by late adolescence. Similarly, Van den Akker et al. (2014), as well as a large cross-cultural study (De Bolle et al., 2015), found the emergence of a gender difference in Neuroticism during adolescence.

Although a growing number of studies have examined age and gender differences in personality traits during late childhood and adolescence, many fewer have tested for such differences during early and middle childhood. The available evidence tentatively suggests that, across these earlier developmental periods, mean levels of Extraversion, Openness, and Activity may decline, girls may already show higher levels of Agreeableness and Conscientiousness than boys, and boys may already show higher levels of Activity than girls (De Fruyt et al., 2006; Eaton, 1994; Eaton & Enns, 1986; Lamb, Chuang, Wessels, Broberg, & Hwang, 2002; Prinzie & Deković, 2008; Slobodskaya & Akhmetova, 2010; Van den Akker et al., 2014). However, results have often been inconsistent across these studies, and more evidence is clearly needed.

Therefore, the present research's second major goal was to examine age and gender differences in the Little Six year by year from early childhood into early adulthood. We were particularly interested in (a) testing for positive (supporting the maturity hypothesis) or U-shaped (supporting the disruption hypothesis) age trends in Conscientiousness, Agreeableness, and Openness during the transition from childhood to adolescence; (b) testing for gender-specific age trends in Neuroticism during these same years; and (c) exploring possible age and gender differences during early and middle childhood.

Looking Beneath the Little Six: Developmental Trends in Nuance Traits

Personality traits can be conceptualized hierarchically, with broader, higher-order traits subsuming narrower, lower-order ones (Markon, Krueger, & Watson, 2005). In the terminology developed by Costa and McCrae (2010; McCrae, *in press*), broad personality "domains," such as Extraversion, each subsume a number of more specific "facet" traits, such as assertiveness and sociability. Each facet, in turn, subsumes "nuance" traits that are narrow enough to be represented by individual

questionnaire items. Pairs of same-domain facet or nuance traits are conceptually and empirically related to each other. However, such traits can also be meaningfully distinguished, with each facet and nuance capturing unique information (e.g., McCrae, *in press*; Paunonen & Ashton, 2001). Moreover, previous studies have found that facet and nuance traits sometimes show distinctive age trends that differ from their superordinate domain (e.g., Lucas & Donnellan, 2009; Roberts et al., 2006; Soto & John, 2012; Soto et al., 2011; Terraciano, McCrae, Brant, & Costa, 2005). These adult findings suggest that lower-order traits may also show distinctive developmental trends in childhood and adolescence. However, few studies have investigated this possibility, and these studies have not converged on a particular pattern of findings (De Fruyt et al., 2006; McCrae et al., 2002; Prinzie & Deković, 2008; Slobodskaya & Akhmetova, 2010).

Therefore, the present research's third major goal was to examine the mean-level development of more specific traits within each broad Little Six dimension. We chose to pursue this goal at the nuance level (using individual CCQ items) rather than the facet level (using multiple-item facet scales) for three reasons. First, there is not yet consensus regarding the most important facet-level youth traits (cf. Costa & McCrae, 2010; Halverson et al., 2003; Mervielde & De Fruyt, 2002). Second, the lower-order structure of youth traits appears to shift with age (Caspi et al., 2005; Soto & John, 2014). Third, the CCQ was developed to minimize conceptual redundancy across items and thereby promote analysis and interpretation of individual items (e.g., Block & Block, 2006). Thus, item-level analysis of the CCQ would allow us to investigate the development of nuance traits using a highly sensitive, bottom-up approach. Specifically, such analyses could identify individual items that show distinctive developmental trends—as well as clusters of items that show trends similar to each other—without imposing a static facet-level structure that may be inappropriate during some developmental periods.

Overview of the Present Research

In sum, the present research was conducted to address three key research questions. First, can the common-language CCQ be used to measure the Little Six? Due to the breadth and depth of the CCQ item pool, we expected that it would be possible to construct a reliable scale for each Little Six dimension. Second, how do mean levels of the Little Six differ by age and gender across childhood, adolescence, and early adulthood? We tentatively expected to find (a) U-shaped age trends for Agreeableness, Conscientiousness, and Openness; (b) negative age trends for Extraversion and Activity; (c) gender differences in Agreeableness and Conscientiousness (with girls scoring higher than boys), as well as Activity (with boys scoring higher than girls), by middle childhood; and (d) gender-specific age trends for Neuroticism, with mean levels inclining across adolescence among girls but not boys. Third, do some nuance traits show developmental trends distinct from their superordinate Little Six

domain? We broadly expected to find some distinctive trends, but we did not have clear predictions regarding specific nuances. We tested these hypotheses by analyzing parent reports for a cross-sectional sample of 16,000 target children between the ages of 3 and 20 years old.

METHOD

Participants and Procedure

Participants were the parents or guardians of 16,000 children, adolescents, and young adults between the ages of 3 and 20 years old. This sample of target children was selected from an initial set of 24,373 to balance for age and gender. Specifically, the SAMPLE command in SPSS 21 was used to randomly select a final sample including 500 males and 500 females in each of 16 age groups: each individual year of age from 3 to 17, plus a combined 18–20-year-old group. In terms of ethnicity, 78% of the target children were described as White/Caucasian, 4% as Black/African American, 4% as Hispanic/Latino, 3% as Asian/Asian American, 1% as Native American/American Indian, 8% as mixed ethnicity, and 2% as another ethnicity. Approximately 83% resided in the United States, 7% in the United Kingdom or Ireland, 6% in Canada, and 4% in Australia or New Zealand. Most of the parents (89%) were mothers.³

Participants anonymously completed a questionnaire describing their child's personality. This questionnaire was hosted on a noncommercial Web site (personalitylab.org) that potential participants could find through search engines, links from other Web sites, or word of mouth. After completing the questionnaire, participants received automatically generated feedback about their child's personality, as well as general information about personality research.

Measures

Common-Language California Child Q-Set. Participants completed a version of the common-language California Child Q-Set (CCQ). The original CCQ (Block & Block, 1980) was developed to allow researchers and clinicians to comprehensively rate youths' personal characteristics. The common-language CCQ (Caspi et al., 1992) revised many of the original items using simpler, nontechnical language so that the measure could be used with parents and other nonprofessional observers. In the present research, we modified the common-language CCQ in two ways. First, we replaced specific words or phrases in 12 items so that they could be applied to adolescents and young adults as well as children. Second, participants independently rated each item on a scale ranging from 1 (*extremely uncharacteristic*) to 9 (*extremely characteristic*), rather than sorting the items into a fixed Q-sort distribution (see Block & Block, 1980). Prior to analysis, we controlled for individual differences in acquiescent responding—the tendency of a respondent to consistently agree or consistently disagree with items, regardless of their content—by centering each participant's set

of ratings around his or her mean score on an acquiescence index that included 26 pairs of opposite items (see Soto & John, 2014).

Development of the CCQ–Little Six Scales. To develop scales for measuring the Little Six, we used a joint rational-empirical approach that drew on previous research using the CCQ. Specifically, we assigned each CCQ item to a Little Six scale if it met two or more of the following criteria: (a) it was rationally classified into the corresponding Big Five dimension by John et al. (1994), (b) it loaded substantially on the corresponding principal component in the present sample (Soto & John, 2014), (c) it loaded on the corresponding component in John et al. (1994), and (d) it loaded on the corresponding component in Van Lieshout and Haselager (1994). We considered a loading substantial if it was at least .40 in strength, or at least .50 in the case of Agreeableness (due to the abundance of Agreeableness content on the CCQ). These criteria thus assigned items based on convergence between rational judgments of item content and previous empirical findings. One CCQ item (“76. Can be trusted; is reliable and dependable”) met the assignment criteria for both Agreeableness and Conscientiousness, and one (“64. Is calm and relaxed; easy-going”) met the criteria for both Agreeableness and (low) Neuroticism. Based on these items' content, we assigned them to the Conscientiousness and Neuroticism scales, respectively. The resulting CCQ–Little Six scales included a total of 67 items, listed in the appendix.

Examination of Measurement Invariance. We conducted a series of multiple-group confirmatory factor analyses (CFAs) to test for scalar invariance of the CCQ–Little Six scales across age and gender (i.e., equality of items' factor loadings and intercepts; Meredith, 1993). Establishing such invariance would indicate that the CCQ–Little Six scales function similarly in different groups, thereby allowing the straightforward interpretation of observed mean-level differences. Conversely, failure to establish invariance would indicate differential item functioning (i.e., one or more items' measurement characteristics differing across groups, relative to the rest of the scale; Reise, Widaman, & Pugh, 1993). Such a result would highlight the importance of examining item-level nuance traits alongside Little Six scale scores.

For each CCQ–Little Six scale, we conducted two pairs of multiple-group CFAs using Mplus 7 (Muthén & Muthén, 2012). The first pair compared a model in which each item's loading and intercept were estimated separately for boys versus girls to a model in which these parameters were constrained to be equal across gender. Their results supported scalar invariance: From the freely estimated to the constrained model, fit statistics that prioritize parsimony by strongly penalizing model complexity indicated increases in fit (e.g., increases in TLI of up to .09; decreases in RMSEA of up to .03), whereas fit statistics with little or no penalty for complexity indicated only trivial decreases in fit (e.g., differences in CFI and SRMR of .01 or less).

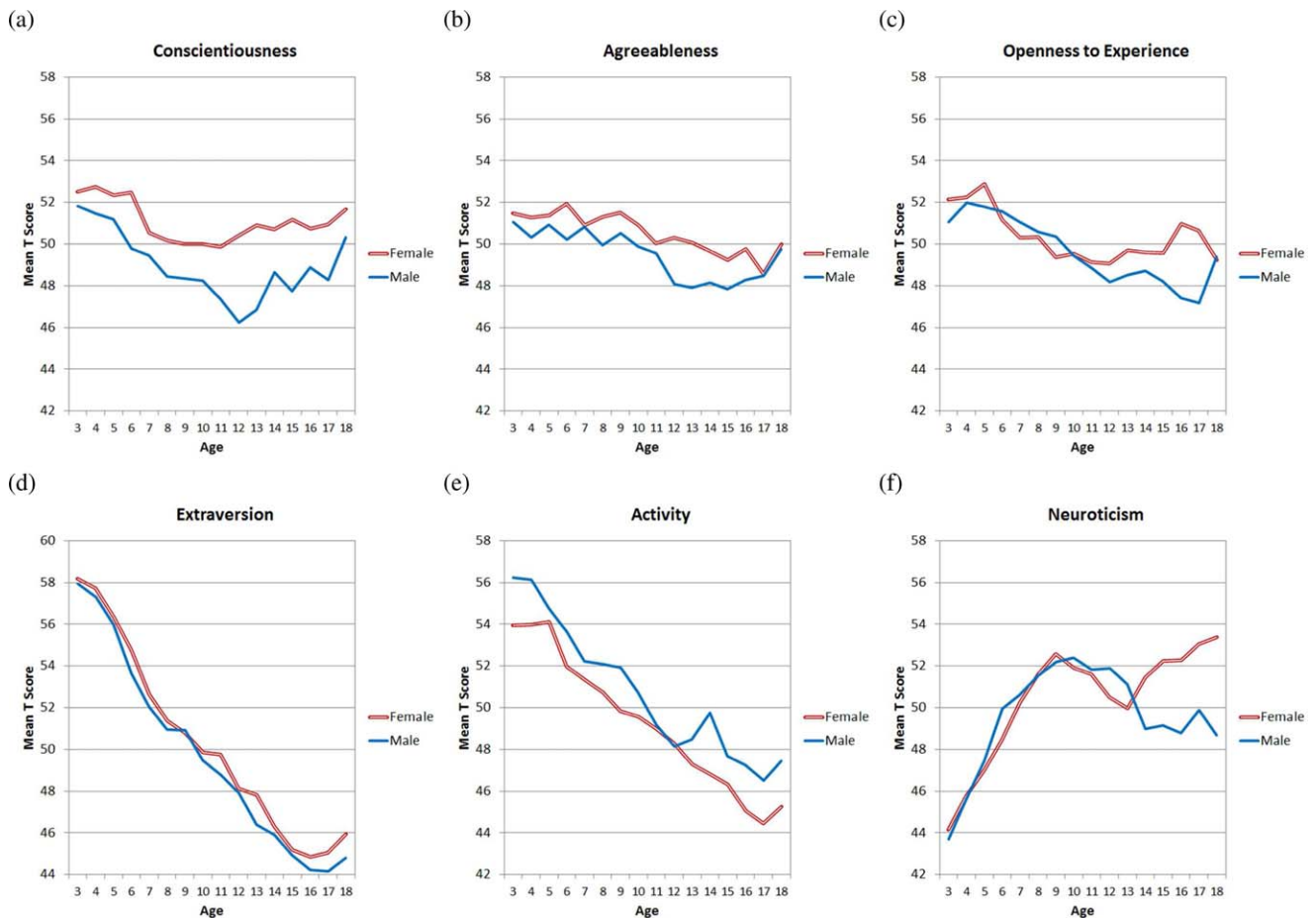


Figure 1 Mean levels of the Little Six by age and gender.

The second pair of CFAs similarly compared a model in which each item's loading and intercept were estimated separately for each of the 16 age groups to a model in which these parameters were constrained to be equal across age. Supporting scalar invariance, from the freely estimated to the constrained model, fit statistics that prioritize parsimony generally indicated increases in fit (e.g., increases in TLI of up to .07; decreases in RMSEA of up to .02). However, supporting differential item functioning, other fit statistics suggested modest to substantial decreases in fit (e.g., decreases in CFI of up to .14; increases in SRMR of up to .08). Taken together, these results indicate that mean-level gender differences in the CCQ–Little Six scales may be interpreted straightforwardly, but that mean-level age differences should be interpreted cautiously and accompanied by item-level analyses of nuance traits.⁴

Data Analysis

We examined mean-level age and gender differences in personality traits in three steps. First, we fit a regression model to identify the overall developmental trend for each dimension. These

models included the effects of gender (coded 1 = female, –1 = male), age (centered at 10.5), age², and age³, as well as interaction terms (see Table 2). Second, to examine these trends in greater detail, we plotted each Little Six dimension's pattern of means by gender and year of age (see Figure 1). Finally, we conducted item-level analyses to examine whether any nuance traits showed distinctive developmental trends. Specifically, we correlated each CCQ–Little Six item's set of 16 age-specific means with the set of means computed from the rest of its scale. A near-zero or negative correlation would indicate that a particular nuance's developmental trend was distinct from the rest of its Little Six dimension. We then plotted and visually inspected the item-specific means to verify interpretation of these correlations.

Due to the large size of the present sample and the exploratory nature of some analyses, we used a conservative significance level of .005 for all hypothesis tests. To facilitate interpretation of results, we scaled the Little Six dimensions and item-level nuances as *T*-scores with a mean of 50 and standard deviation of 10 while controlling for between-group variability (cf. Soto et al., 2011). In terms of Cohen's (1988) guidelines for effect size, a group difference of 2 *T*-points indicates a small effect, a difference of 5 *T*-points indicates a medium-sized

Table 1 Descriptive Statistics, Reliability Coefficients, and Correlations of the CCQ–Little Six Scales

	Items	M	SD	α	Correlations					
					Ext.	Agr.	Con.	Neu.	Ope.	Act.
CCQ–Little Six Scales										
Extraversion	9	0.42	1.70	.83	—	—	—	—	—	—
Agreeableness	22	0.53	1.53	.93	.09	—	—	—	—	—
Conscientiousness	10	0.23	1.52	.86	.10	.63	—	—	—	—
Neuroticism	15	–0.23	1.43	.87	–.41	–.32	–.45	—	—	—
Openness to Experience	6	1.37	1.21	.63	.13	.20	.25	–.14	—	—
Activity	5	0.66	1.67	.73	.35	.16	.33	–.42	.11	—
Principal Components										
Extraversion	—	—	—	—	.89	.05	.06	–.19	.12	.13
Agreeableness	—	—	—	—	.03	.91	.36	–.19	.19	.10
Conscientiousness	—	—	—	—	–.10	.34	.76	–.10	–.04	–.05
Neuroticism	—	—	—	—	–.27	–.11	–.25	.91	–.02	–.25
Openness to Experience	—	—	—	—	.04	.07	.21	–.03	.87	.04
Activity	—	—	—	—	.13	.05	.35	–.18	.04	.84

Note. $N = 16,000$. CCQ–Little Six scales are scored from within-person centered items. Principal components are extracted from a set of 94 personality-relevant CCQ items and rotated using the varimax criterion (Soto & John, 2014). All correlations are statistically significant at $p < .005$. α = Cronbach's alpha reliability coefficient; Ext. = Extraversion; Agr. = Agreeableness; Con. = Conscientiousness; Neu. = Neuroticism; Ope = Openness to Experience; Act. = Activity.

Table 2 Regression Coefficients Predicting the Little Six From Age and Gender

	Conscientiousness	Agreeableness	Openness to Experience	Extraversion	Activity	Neuroticism
	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)
Intercept	48.9373 (.1190)*	49.9013 (.1190)*	49.4975 (.1191)*	49.0487 (.1190)*	49.5910 (.1191)*	51.6464 (.1191)*
Gender	1.3281 (.1190)*	.7025 (.1190)*	.1168 (.1191)	.2584 (.1190)	–.5143 (.1191)*	–.1830 (.1191)
Age	–.1834 (.0435)*	–.3759 (.0435)*	–.3335 (.0435)*	–1.0230 (.0435)*	–.7140 (.0435)*	–.1121 (.0435)
Gender \times Age	.1326 (.0435)*	.0733 (.0435)	.2115 (.0435)*	–.0076 (.0435)	–.0318 (.0435)	.1317 (.0435)*
Age ²	.0500 (.0042)*	.0046 (.0042)	.0236 (.0042)*	.0448 (.0042)*	.0192 (.0042)*	–.0775 (.0042)*
Gender \times Age ²	–.0124 (.0042)*	–.0086 (.0042)	.0115 (.0042)	.0015 (.0042)	–.0114 (.0042)	.0270 (.0042)*
Age ³	.0012 (.0011)	.0051 (.0011)*	.0028 (.0011)	.0021 (.0011)	.0014 (.0011)	.0115 (.0011)*
Gender \times Age ³	–.0021 (.0011)	–.0017 (.0011)	–.0038 (.0011)*	.0006 (.0011)	.0005 (.0011)	.0003 (.0011)

Note. $N = 16,000$. Gender is coded 1 = female, –1 = male. Age is measured in years and centered at 10.5. The CCQ–Little Six are scaled as T -scores ($M = 50$, $SD = 10$).

* $p < .005$.

effect, and a difference of 8 T -points or greater indicates a large effect. For pairwise comparisons with 500 observations per group (e.g., comparing two different age groups within the same gender, or comparing boys vs. girls within the same age group), mean-level differences of 2 T -points or greater are statistically significant at the .005 level. Note that all effects reported here should be interpreted as cross-sectional differences rather than longitudinal changes.

RESULTS

The CCQ–Little Six Scales

Table 1 presents descriptive statistics, alpha reliability coefficients, and intercorrelations for the CCQ–Little Six scales, as well as these scales' correlations with varimax-rotated principal

components derived from a set of 94 personality-relevant CCQ items (Soto & John, 2014). Reliability coefficients were quite high for the Agreeableness (22 items, $\alpha = .93$), Neuroticism (15 items, $\alpha = .87$), Conscientiousness (10 items, $\alpha = .86$), and Extraversion (nine items, $\alpha = .83$) scales, and acceptable for the shorter Activity (five items, $\alpha = .73$) and Openness (six items, $\alpha = .63$) scales. Intercorrelations among the scales averaged .27 in strength, and most were weak or moderate. The most notable exception to this pattern was the substantial intercorrelation between Agreeableness and Conscientiousness ($r = .63$). We suspect that this correlation reflects (a) the CCQ's ample representation of obedience and responsibility content (two lower-order traits related to both Agreeableness and Conscientiousness), (b) the general tendency of parents to describe children in terms of a broad easy versus difficult dimension (De Pauw et al., 2009; Tackett et al., 2012), and (c) the possibility that agreeable

and conscientious behaviors may in fact covary more strongly in childhood than in adulthood (Mervielde & De Fruyt, 2002; Soto, John, Gosling, & Potter, 2008). We attempted to reduce the correlation between these two scales through item deletion, but we found that even much shorter and conceptually narrower versions remained substantially intercorrelated. Based on this finding, and the importance of maintaining appropriate breadth of content when measuring broad constructs, we retained the full Agreeableness and Conscientiousness scales reported in the appendix.

As shown in the bottom half of Table 1, convergent correlations of the CCQ–Little Six scales with varimax-rotated principal components were uniformly strong: They averaged .86, and all were at least .76. In contrast, the discriminant correlations averaged only .14 in strength, and all were weaker than .40. Taken together, the results reported in Table 1 indicate that the CCQ can be used to reliably measure each Little Six dimension.

Age and Gender Differences in the Little Six and Nuance Traits

Conscientiousness

Coefficients from analyses regressing Conscientiousness on age and gender are presented in Table 2, and mean levels of Conscientiousness by age and gender are shown in Figure 1a. Overall Conscientiousness showed a quadratic age trend ($b = .0500$, $p < .005$), which was more pronounced among boys than girls, as indicated by the significant $\text{Gender} \times \text{Age}^2$ interaction term ($b = -.0124$, $p < .005$). Specifically, mean levels declined from early to late childhood by 6 T -points among boys and 3 T -points among girls and then inclined from early adolescence into early adulthood by 4 T -points among boys and 2 T -points among girls. Girls were more conscientious than boys at every age, with this gender difference largest (approximately 4 T -points) during early adolescence.

Analyses of nuance traits indicated that 8 of the 10 Conscientiousness items showed age trends similar to overall Conscientiousness. One exception was “99. Thinks about their actions and behavior; uses their head before doing or saying something,” which showed a flat age trend from early to middle childhood and then the normative incline from early adolescence into early adulthood. The second exception was “47. Has high standards for themselves; needs to do very well in the things they do,” which inclined from early to middle childhood by 2 T -points, then showed the normative U-shaped trend. Taken together, these results support the disruption hypothesis for most nuances of Conscientiousness. They further indicate that girls tend to be more conscientious than boys, especially during late childhood and early adolescence.

Agreeableness. Regression coefficients for Agreeableness are presented in Table 2, and mean levels are shown in Figure 1b.

Overall Agreeableness showed a cubic age trend ($b = .0051$, $p < .005$). Mean levels were consistent across early and middle childhood, declined by approximately 3 T -points from late childhood through early adolescence, and then inclined by approximately 2 T -points from late adolescence into early adulthood. Girls tended to be slightly more agreeable than boys, but these gender differences were small (2 T -points or fewer) at every age, and the age trend for Agreeableness was similar among boys and girls.

Nuance-level analyses indicated that several Agreeableness items showed age trends distinct from overall Agreeableness. These distinctive items fell into three clusters. First, the items “31. Is able to see how others feel; can put themselves in another person’s place,” “13R. Tries to see what and how much they can get away with; usually pushes limits and tries to stretch rules,” and “93R. Is bossy and likes to dominate other people” showed monotonic age trends that indicated greater perspective taking, less bossiness, and less limit pushing at older ages, with total age differences of 2–7 T -points. Second, the items “11R. Tries to blame other people for things they have done themselves,” “55R. Worries about not getting their share of material things, food, or love; seems afraid they won’t get enough,” and “95R. Lets little problems get to them and is easily upset. It doesn’t take much to get them irritated or mad” showed roughly quadratic age trends. These trends indicated inclines (of 2–7 T -points) in blaming others, worrying about fairness, and irritability from early to late childhood and then declines (of 2–4 T -points) from early adolescence into early adulthood. The final cluster included “62. Is obedient and does what they are told,” “22R. Tries to get others to do things by playing up to them; acts charming in order to get their way,” “85R. Is aggressive. (For example, picks fights or starts arguments),” and “90R. Is stubborn.” Their age trends indicated inclines in obedience and declines in manipulation, aggression, and stubbornness across early and middle childhood (with totals of 2–5 T -points), followed by the normative U-shaped pattern from late childhood into early adulthood. These results support the disruption hypothesis for overall Agreeableness and indicate that girls tend to be somewhat more agreeable than boys. However, they also indicate that many nuances of Agreeableness show distinctive developmental trends.

Openness to Experience. Regression coefficients for Openness are presented in Table 2, and mean levels are shown in Figure 1c. As indicated by the significant $\text{Gender} \times \text{Age}^3$ interaction term ($b = -.0038$, $p < .005$), the age trend for overall Openness differed by gender. Among both boys and girls, mean levels were consistent from early to middle childhood and then showed a roughly U-shaped age trend. However, the precise shape of this trend differed between boys and girls. Among boys, Openness declined by 5 T -points from middle childhood into late adolescence and then inclined by 2 T -points from late adolescence into early adulthood. Among girls, Openness declined by 4 T -points from middle to late childhood, inclined by 2 T -points from early to late adolescence, and then declined

by 2 *T*-points from late adolescence into early adulthood. As a result of these trends, girls tended to be somewhat more open than boys during adolescence, with gender differences of up to 3 *T*-points.

Nuance-level analyses indicated that three of the six Openness items showed distinctive age trends. One distinctive item was “40. Is curious and exploring; likes to learn and experience new things,” which declined across childhood and early adolescence by approximately 10 *T*-points and was then consistent into early adulthood. The second item was “70. Daydreams; often gets lost in thought or a fantasy world.” This item inclined across childhood and adolescence by approximately 6 *T*-points; from late adolescence into early adulthood, it showed a flat age trend for boys but declined by 2 *T*-points for girls. The final item was “68. Is a very smart person (even though formal tests and evaluations might not show this).” This item very gradually inclined with age among boys, by a total of 2 *T*-points, and remained flat among girls. These results support the disruption hypothesis for overall Openness. However, they also indicate that Openness’s age trend differs somewhat for boys versus girls, and that two important nuance traits—curiosity and absorption in fantasy—show distinctive and virtually opposite trends.

Extraversion. Regression coefficients for Extraversion are presented in Table 2, and mean levels are shown in Figure 1d. Overall Extraversion showed a quadratic age trend ($b = .0448$, $p < .005$). Mean levels declined from early childhood through early adolescence by approximately 13 *T*-points and then remained consistent into early adulthood. This pattern was very similar for boys and girls, and there were no more than trivial gender differences at any age. Nuance-level analyses indicated that all nine Extraversion items showed age trends similar to overall Extraversion. These results support the hypothesis that youths’ expressive behavior declines steadily across childhood and early adolescence.

Activity. Regression coefficients for Activity are presented in Table 2, and mean levels are shown in Figure 1e. Overall Activity showed a quadratic age trend ($b = .0192$, $p < .005$). Mean levels declined from early childhood into late adolescence by approximately 10 *T*-points and then remained consistent into early adulthood. Boys tended to be somewhat more active than girls, with gender differences of up to 3 *T*-points.

Nuance-level analyses indicated that two of the five Activity items showed distinctive age trends. One such item was “37. Likes to compete; is always testing and comparing himself to other people,” which inclined from early to middle childhood by approximately 3 *T*-points and then declined from late childhood into early adulthood by approximately 3 *T*-points. The other item was “51. Is well-coordinated. (For example, does well in sports).” Among girls, this item steadily declined, by a total of 4 *T*-points. Among boys, it declined from early to middle childhood by 4 *T*-points and then inclined from late childhood into early adulthood by 2 *T*-points. These results support the hypoth-

eses that physical energy level declines substantially across childhood and adolescence, and that boys tend to be more active than girls from a young age. They also indicate that competitive drive shows a distinctive developmental trend.

Neuroticism. Regression coefficients for Neuroticism are presented in Table 2, and mean levels are shown in Figure 1f. Overall Neuroticism showed a cubic age trend ($b = .0115$, $p < .005$). For both boys and girls, mean levels inclined by approximately 9 *T*-points from early to middle childhood. However, as indicated by the significant Gender \times Age ($b = .1317$, $p < .005$) and Gender \times Age² ($b = .0270$, $p < .005$) interaction terms, Neuroticism showed quite different age trends for boys versus girls from late childhood into early adulthood. Among boys, it declined by 4 *T*-points from late childhood into early adulthood. Among girls, it declined by 2 *T*-points across late childhood, but then inclined by 3 *T*-points from early adolescence into early adulthood. As a result of these trends, a gender difference in Neuroticism emerged during adolescence, with girls scoring approximately 5 *T*-points higher than boys by early adulthood.

Nuance-level analyses indicated that 12 of the 15 Neuroticism items showed age trends similar to overall Neuroticism. One exception was “77. Feels unworthy; has a low opinion of themselves.” This item inclined monotonically across childhood and adolescence, with a total incline of approximately 13 *T*-points. The second exception was “33. Cries easily,” which was consistent across early and middle childhood and then declined from late childhood into early adulthood by 4 *T*-points for girls and 10 *T*-points for boys. The final exception was “64R. Is calm and relaxed; easy-going.” This item indicated consistent levels of calmness across early and middle childhood and then gender-specific age trends consistent with overall Neuroticism from late childhood into early adulthood. These results support the hypothesis that Neuroticism shows gender-specific age trends during adolescence, leading to greater negative emotionality among girls than boys. They further indicate that, among both boys and girls, mean levels of overall Neuroticism incline during childhood, self-doubt inclines throughout childhood and adolescence, and crying declines during adolescence.

DISCUSSION

Measuring the Little Six Youth Personality Dimensions

The present research pursued three main goals. The first was to develop a method for measuring the Little Six youth personality dimensions—Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Activity—using the item pool of the California Child Q-set. We used a joint rational-empirical approach to construct the CCQ–Little Six scales and found that they reliably assess each Little Six dimension, including assessment of Activity as largely independent from the Big Five.

Conceptually, the Little Six represent a union of the most prominent trait dimensions from the personality and temperament literatures, and they can therefore help to integrate these two research traditions within a common organizing framework (De Pauw & Mervielde, 2010; Shiner & DeYoung, 2013; Soto & John, 2014). Empirically, several studies support the Little Six as a structural model of youths' traits (De Pauw et al., 2009; John et al., 1994; Soto & John, 2014; Van Lieshout & Haselager, 1994). To our knowledge, however, the CCQ–Little Six is the first measure that independently assesses each Little Six dimension. This measure should therefore prove useful for future studies examining youths' psychological characteristics: studies that analyze the existing archive of CCQ data, as well as studies that collect new data. The CCQ–Little Six can also inform the development of future Little Six measures. For example, a new measure could aim to include a more even balance of item content (e.g., more Openness and Activity content) and better differentiate between some Little Six dimensions (e.g., Agreeableness and Conscientiousness).

Mean-Level Age and Gender Differences in the Little Six

Our second major goal was to examine mean-level age and gender differences in the Little Six across childhood, adolescence, and early adulthood. We found a complex pattern of developmental trends. Conscientiousness, Agreeableness, and Openness all showed curvilinear, U-shaped patterns, with declines from childhood into early adolescence followed by inclines from late adolescence into early adulthood. In contrast, Extraversion and Activity steadily declined with age. We also found gender differences in some traits. Boys were consistently more active than girls, whereas girls tended to be more conscientious and agreeable than boys. Neuroticism showed gender-specific age trends: Mean levels inclined across childhood for both boys and girls, and then a gender difference emerged during adolescence as Neuroticism further inclined among girls but declined among boys.

Many of these findings are consistent with results from previous cross-sectional and longitudinal studies (e.g., De Bolle et al., 2015; Denissen et al., 2013; Eaton, 1994; Eaton & Enns, 1986; Lamb et al., 2002; Prinzie & Deković, 2008; Slobodskaya & Akhmetova, 2010; Soto et al., 2011; Van den Akker et al., 2014). There is thus growing evidence that many personality traits show substantial mean-level age and gender differences across childhood and adolescence. Interestingly, this evidence indicates that Conscientiousness, Agreeableness, and Openness show patterns consistent with the disruption hypothesis: Rather than steadily increasing throughout the life span, some aspects of psychosocial maturity temporarily decline during the transition from childhood to adolescence (cf. Denissen et al., 2013; Soto et al., 2011; Van den Akker et al., 2014). There is also growing evidence that boys and girls show somewhat different patterns of personality development (cf. De Bolle et al., 2015;

Eaton & Enns, 1986; Slobodskaya & Akhmetova, 2010; Soto et al., 2011; Van den Akker et al., 2014), with early appearing and persistent gender differences in some traits (especially Activity and Conscientiousness) and gender-specific age trends for others (especially Neuroticism). All of these findings highlight the importance of additional research examining youth personality development.

Mean-Level Development of Personality Nuances

Our final goal was to examine the development of Little Six nuances, as captured by individual CCQ items. Many nuance traits showed distinctive trends. For example, perspective taking, self-doubt, and absorption in fantasy steadily inclined with age, whereas attention seeking, bossiness, crying, and curiosity steadily declined, all of which contrast with the trends shown by overall Agreeableness, Neuroticism, and Openness. We also found some distinctive nuance-level trends specific to early and middle childhood. Across these years, competitiveness and self-standards inclined, despite declines in overall Activity and Conscientiousness.

These findings agree with results from previous studies in suggesting that—throughout the life span—the more specific traits within a broad personality dimension sometimes show distinctive developmental trends (e.g., Roberts et al., 2006; Soto & John, 2012; Soto et al., 2011; Terracciano et al., 2005; Van den Akker et al., 2014). Moreover, the present results illustrate that such trends can be identified through nuance-level analyses of individual items (see also Lucas & Donnellan, 2009; McCrae, in press). Nuance-level analyses offer a highly sensitive, bottom-up approach: They can identify clusters of individual items with similar developmental trends, without imposing a static facet-level structure. This approach may be particularly appropriate in childhood and adolescence, when the lower-order structure of youths' personality traits is less clear and may shift with age (Caspi et al., 2005; Soto & John, 2014).

In fact, nuance-level trends may help explain some differences in youth versus adult personality structure. For example, within the Activity dimension, we found marked declines in physical activity level—but not competitive drive—with age. These trends may reflect the process by which Activity recedes from a major personality dimension in childhood to a lower-order trait in adulthood. In childhood, Activity is primarily defined by physical energy and motor activity (Soto & John, 2014). These physical nuances are prominent features of many children's personalities, as indicated by their high mean levels (Figure 1e; see also Eaton, 1994). By adolescence, the meaning of Activity expands to include psychological nuances such as motivation and competitive drive (Soto & John, 2014), whereas mean levels of physical activity decline. Physical activity level declines still further into young adulthood, such that overall Activity no longer appears prominent enough to qualify as a major personality dimension (Eaton, 1994). Instead, Activity's

physical nuances become integrated into Big Five Extraversion, and its motivational nuances into Conscientiousness (Costa & McCrae, 2010; Shiner & DeYoung, 2013). These patterns highlight the importance of assessing Activity in studies of youth personality, either as a major, independent dimension (especially during childhood and early adolescence) or as a facet-level trait (especially during late adolescence and young adulthood).

LIMITATIONS AND FUTURE DIRECTIONS

The present research had a number of important strengths, including its large sample, broad age range, independent assessment of each Little Six dimension, and examination of nuance traits. However, it also had some important limitations. For example, all of the present participants volunteered to complete an online survey; therefore, the present sample may not fully represent the general population. Moreover, the present data were all parent reports. An important advantage of parent reports, compared with alternatives such as teacher reports and self-reports, is that they allowed us to assess personality traits across a broad age range spanning from early childhood (when most children have not started school and cannot provide valid self-reports) into early adulthood. However, parent reports are imperfect indicators of personality that sometimes disagree with other sources of information (e.g., De Fruyt et al., 2006; Van den Akker et al., 2014). In particular, parent reports tend to be positively biased, and they may be less accurate for adolescents and young adults who no longer live at home (including, presumably, many target children in the present 18–20-year-old age group). They may also be shaped by parental expectations regarding group differences in youths' behavior; for example, parents may expect girls to be less physically active than boys and thus rate girls lower on Activity. Another limitation of our measurement approach was that using individual items to assess nuance traits restricts the reliability of these measurements (McCrae, in press). A final noteworthy limitation was the present research's cross-sectional design. Cross-sectional age differences may reflect not only true developmental changes, but also cohort effects (i.e., the effects of different children being born and raised in different historical contexts) and differential selection effects (i.e., the effects of parents choosing to participate for different reasons at different ages).

Many of these concerns are partially alleviated by convergence between the present findings and previous studies that have used other recruitment strategies (e.g., Slobodskaya & Akhmetova, 2010), other measurement approaches (e.g., Soto et al., 2011), and longitudinal or meta-analytic designs (e.g., Denissen et al., 2013; Van den Akker et al., 2014). However, additional replication is still needed. Further evidence regarding personality development prior to age 10, as well as the development of nuance traits, will be particularly valuable.

Beyond these limitations, the present research raises important questions. For example, to what extent do the developmental trends observed in studies of parent reports, teacher reports, and youths' self-reports reflect age and gender differences in youths' behavior, and to what extent do they reflect differences in the standards that raters use to make personality judgments (see also Denissen et al., 2013)? To some degree, parents in the present sample appear to be holding children of different ages to different standards. For example, as described above, the item "68. Is a very smart person (even though formal tests and evaluations might not show this)" shows only a small mean-level incline with age. This clearly does not indicate that a typical 3-year-old is almost as smart as a typical 18-year-old; a high school senior obviously possesses many more cognitive skills than a preschooler. Instead, parents are likely comparing 3-year-olds and 18-year-olds to different standards of cognitive ability.

Despite such shifts in rating standards, there are at least three important reasons to believe that most observed age and gender differences in personality ratings reflect differences in actual behavior. First, in the present study, the CCQ–Little Six scales showed a reasonable degree of measurement invariance across age and gender, suggesting that they function similarly (although not identically) in different groups. Second, if parents, teachers, and youths accurately shift their rating standards with age, then these shifts will tend to diminish rather than magnify age differences in personality ratings. Therefore, age differences in personality ratings are more likely to represent conservative estimates than overestimates of true age differences in personality. Third, many observed age and gender differences in personality ratings converge with other sources of information. For example, U-shaped age trends in Agreeableness and Conscientiousness converge with peaks in risky, oppositional, and antisocial behavior during adolescence (Centers for Disease Control and Prevention, 2012; Federal Bureau of Investigation, 2013; Moffitt, 1993; Steinberg et al., 2008), and gender differences in Conscientiousness parallel differences in academic achievement (Fortin, Oreopoulos, & Phipps, 2013). Age and gender differences in Activity converge with studies using objective instruments, such as actometers and pedometers (Eaton, 1994; Eaton & Enns, 1986). Gender-specific trends for Neuroticism during adolescence mirror the emergence of gender differences in clinical anxiety and depression (Lewinsohn, Gotlib, Lewinsohn, Seeley, & Allen, 1998; Nolen-Hoeksema & Girgus, 1994). However, additional research is needed to further disentangle gender and age differences in youths' behavior from differences in rating standards. Such research could benefit from obtaining measures of personality and behavior based on direct observation, as well as traditional self-reports and informant reports (e.g., Borkenau, Reimann, Angleitner, & Spinath, 2001).

A final, crucial question concerns causality. What developmental mechanisms underlie age and gender differences in youths' personality traits? Previous research investigating the origins of personality differences (e.g., Borkenau et al., 2001; Jang, Livesley, & Vernon, 1996), as well as predictors of adult personality change (e.g., Bleidorn et al., 2013; McCrae et al.,

2000; Specht et al., 2014), indicate that personality development is influenced by both biological and social factors. Biological influences on youth personality development likely include genetics, as well as changes in hormones, brain structure, and brain chemistry (Bleidorn, Kandler, Riemann, Angleitner, & Spinath, 2009; DeYoung & Gray, 2009). Social causes likely include family, peer, and academic experiences (Bleidorn, 2012; Shiner & Caspi, 2003; Van den Akker et al., 2014). However, much additional research—especially longitudinal studies designed to test whether earlier biological and social factors predict subsequent personality change—is needed to identify the specific mechanisms underlying youth personality development. As consensus builds regarding patterns of age and gender differences in youths' personality traits, the shift from describing such differences to explaining them will become increasingly important.

Declaration of Conflicting Interests

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

Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Preparation of this article was supported by a faculty research grant from Colby College.

Notes

1. In the remainder of this article, we use the term *personality* broadly to include traits drawn from both the child temperament and adult personality literatures.
2. Throughout this article, we use the terms *incline* (i.e., positive age trend) and *decline* (i.e., negative age trend) to describe cross-sectional age differences. We reserve the terms *increase* and *decrease* for describing longitudinal changes.
3. The present sample was also analyzed by Soto and John (2014), who examined personality structure but not mean-level age or gender differences in personality traits.
4. An alternative approach would be to eliminate items that show differential functioning. However, this approach would result in narrower definitions of the Little Six; it would also provide less total information than examining each conceptually relevant item individually.

References

- Bleidorn, W. (2012). Hitting the road to adulthood: Short-term personality development during a major life transition. *Personality and Social Psychology Bulletin*, *38*, 1594–1608.
- Bleidorn, W., Kandler, C., Riemann, R., Angleitner, A., & Spinath, F. M. (2009). Patterns and sources of adult personality development: Growth curve analyses of the NEO PI-R scales in a longitudinal twin study. *Journal of Personality and Social Psychology*, *97*, 142–155.
- Bleidorn, W., Klimstra, T. A., Denissen, J. A., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2013). Personality maturation around the world: A cross-cultural examination of social-investment theory. *Psychological Science*, *24*, 2530–2540.
- Block, J., & Block, J. H. (1980). *The California Child Q-Set*. Palo Alto, CA: Consulting Psychologists Press.
- Block, J., & Block, J. H. (2006). Nursery school personality and political orientation two decades later. *Journal of Research in Personality*, *40*, 734–749.
- Borkenau, P., Riemann, R., Angleitner, A., & Spinath, F. M. (2001). Genetic and environmental influences on observed personality: Evidence from the German Observational Study of Adult Twins. *Journal of Personality and Social Psychology*, *80*, 655–668.
- Buss, A. H., & Plomin, R. (1984). *A temperament theory of personality development* (2nd ed.). New York: Wiley.
- Caspi, A., Block, J., Block, J. H., Klopp, B., Lynam, D., Moffitt, T. E., et al. (1992). A “common-language” version of the California Child Q-Set for personality assessment. *Psychological Assessment*, *4*, 512–523.
- Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, *56*, 453–484.
- Centers for Disease Control and Prevention. (2012). *Youth risk behavior surveillance, United States, 2011*. Washington, DC: U.S. Department of Health and Human Services.
- Clark, L., & Watson, D. (2008). Temperament: An organizing paradigm for trait psychology. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 265–286). New York: Guilford Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Collins, W. (2003). More than myth: The developmental significance of romantic relationships during adolescence. *Journal of Research on Adolescence*, *13*, 1–24.
- Costa, P. T., & McCrae, R. R. (2010). *NEO Inventories professional manual*. Odessa, FL: Psychological Assessment Resources.
- De Bolle, M., De Fruyt, P., McCrae, R. R., Löckenhoff, C. E., Costa, P. T., Aguilar-Vafaie, M. E., et al. (2015). The emergence of sex differences in personality traits in early adolescence: A cross-sectional, cross-cultural study. *Journal of Personality and Social Psychology*, *108*, 171–185.
- De Fruyt, F., Bartels, M., van Leeuwen, K. G., De Clercq, B., Decuyper, M., & Mervielde, I. (2006). Five types of personality continuity in childhood and adolescence. *Journal of Personality and Social Psychology*, *91*, 538–552.
- Denissen, J. A., Van Aken, M. G., Penke, L., & Wood, D. (2013). Self-regulation underlies temperament and personality: An integrative developmental framework. *Child Development Perspectives*, *7*, 255–260.
- De Pauw, S. W., & Mervielde, I. (2010). Temperament, personality and developmental psychopathology: A review based on the conceptual dimensions underlying childhood traits. *Child Psychiatry and Human Development*, *41*, 313–329.

- De Pauw, S. W., Mervielde, I., & Van Leeuwen, K. G. (2009). How are traits related to problem behavior in preschoolers? Similarities and contrasts between temperament and personality. *Journal of Abnormal Child Psychology*, *37*, 309–325.
- DeYoung, C. G., & Gray, J. R. (2009). Personality neuroscience: Explaining individual differences in affect, behaviour and cognition. In P. J. Corr & G. Matthews (Eds.), *The Cambridge handbook of personality psychology* (pp. 323–346). New York: Cambridge University Press.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*, 1040–1048.
- Eaton, W. O. (1994). Temperament, development, and the five-factor model: Lessons from activity level. In C. F. Halverson, G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 173–187). Hillsdale, NJ: Erlbaum.
- Eaton, W. O., & Enns, L. R. (1986). Sex differences in human motor activity level. *Psychological Bulletin*, *100*, 19–28.
- Erikson, E. H. (1968). *Identity: Youth and crisis*. New York: Norton.
- Federal Bureau of Investigation. (2013). *Uniform crime report for the United States, 2012*. Washington, DC: U.S. Government Printing Office.
- Fortin, N. M., Oreopoulos, P., & Phipps, S. (2013). *Leaving boys behind: Gender disparities in high academic achievement (NBER Working Paper 19331)*. Cambridge, MA: National Bureau of Economic Research.
- Goldberg, L. R. (1990). An alternative “description of personality”: The Big-Five factor structure. *Journal of Personality and Social Psychology*, *59*, 1216–1229.
- Goldsmith, H., Buss, A. H., Plomin, R., Rothbart, M. K., Thomas, A., Chess, S., et al. (1987). What is temperament? Four approaches. *Child Development*, *58*, 505–529.
- Halverson, C. F., Havill, V. L., Deal, J., Baker, S. R., Victor, J. B., Pavlopoulos, V., et al. (2003). Personality structure as derived from parental ratings of free descriptions of children: The Inventory of Child Individual Differences. *Journal of Personality*, *71*, 995–1026.
- Harter, S. (2006). The self. In W. Damon & R. M. Lerner (Series Eds.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 505–570). Hoboken, NJ: Wiley.
- Hill, J. P., & Lynch, M. E. (1983). The intensification of gender-related role expectations during early adolescence. In J. Brooks-Gunn & A. C. Peterson (Eds.), *Girls at puberty* (pp. 201–228). New York: Springer.
- Inhelder, B., & Piaget, J. (1958). *Growth of logical thinking from childhood to adolescence*. New York: Basic Books.
- Jang, K. L., Livesley, W. J., & Vernon, P. A. (1996). The genetic basis of personality at different ages: A cross-sectional twin study. *Personality and Individual Differences*, *21*, 299–301.
- John, O. P., Caspi, A., Robins, R. W., Moffitt, T. E., & Stouthamer-Loeber, M. (1994). The “little five”: Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development*, *65*, 160–178.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114–158). New York: Guilford Press.
- Keating, D. P. (2004). Cognitive and brain development. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 45–84). Hoboken, NJ: Wiley.
- Lamb, M. E., Chuang, S. S., Wessels, H., Broberg, A. G., & Hwang, C. (2002). Emergence and construct validation of the Big Five factors in early childhood: A longitudinal analysis of their ontogeny in Sweden. *Child Development*, *73*, 1517–1524.
- Lewinsohn, P. M., Gotlib, I. H., Lewinsohn, M., Seeley, J. R., & Allen, N. B. (1998). Gender differences in anxiety disorders and anxiety symptoms in adolescents. *Journal of Abnormal Psychology*, *107*, 109–117.
- Lucas, R. E., & Donnellan, M. (2009). Age differences in personality: Evidence from a nationally representative Australian sample. *Developmental Psychology*, *45*, 1353–1363.
- Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology*, *88*, 139–157.
- Marshall, W. A., & Tanner, J. M. (1986). Puberty. In F. Falkner & J. M. Tanner (Eds.), *Human growth* (Vol. *11*, pp. 171–209). New York: Plenum Press.
- McCrae, R. R. (in press). A more nuanced view of reliability: Specificity in the trait hierarchy. *Personality and Social Psychology Review*.
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*, 81–90.
- McCrae, R. R., Costa, P. T., Ostendorf, F., Angleitner, A., Hrebickova, M., & Avia, M. D. (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., Terracciano, A., Parker, W. D., Mills, C. J., De Fruyt, F., et al. (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.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543.
- Mervielde, I., & De Fruyt, F. (2002). Assessing children’s traits with the Hierarchical Personality Inventory for Children. In B. De Raad & M. Perugini (Eds.), *Big Five assessment* (pp. 129–142). Seattle, WA: Hogrefe & Huber.
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674–701.
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus user’s guide* (7th ed.). Los Angeles: Muthén & Muthén.
- Nolen-Hoeksema, S., & Girgus, J. S. (1994). The emergence of gender differences in depression during adolescence. *Psychological Bulletin*, *115*, 424–443.

- Paunonen, S. V., & Ashton, M. C. (2001). Big Five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology, 81*, 524–539.
- Prinzle, P., & Deković, M. (2008). Continuity and change of childhood personality characteristics through the lens of teachers. *Personality and Individual Differences, 45*, 82–88.
- Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114*, 552–566.
- Rice, K. G., & Mulkeen, P. (1995). Relationships with parents and peers: A longitudinal study of adolescent intimacy. *Journal of Adolescent Research, 10*, 338–357.
- 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.
- Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 11–39). Mahwah, NJ: Erlbaum.
- Rothbart, M. K. (2007). Temperament, development, and personality. *Current Directions in Psychological Science, 16*, 207–212.
- Rothbart, M. K., Ahadi, S. A., Hersey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development, 72*, 1394–1408.
- Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology, 94*, 168–182.
- Shiner, R., & Caspi, A. (2003). Personality differences in childhood and adolescence: Measurement, development, and consequences. *Journal of Child Psychology and Psychiatry, 44*, 2–32.
- Shiner, R. L., & DeYoung, C. (2013). The structure of temperament and personality traits: A developmental perspective. In P. D. Zelazo (Ed.), *The Oxford handbook of developmental psychology: Vol. 2. Self and other* (pp. 113–141). New York: Oxford University Press.
- Slobodskaya, H. R., & Akhmetova, O. A. (2010). Personality development and problem behavior in Russian children and adolescents. *International Journal of Behavioral Development, 34*, 441–451.
- Soto, C. J., & John, O. P. (2012). Development of Big Five domains and facets in adulthood: Mean-level age trends and broadly versus narrowly acting mechanisms. *Journal of Personality, 80*, 881–914.
- Soto, C. J., & John, O. P. (2014). Traits in transition: The structure of parent-reported personality traits from early childhood to early adulthood. *Journal of Personality, 82*, 182–199.
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of Big Five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology, 94*, 718–737.
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big-Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology, 100*, 330–348.
- Specht, J., Bleidorn, W., Denissen, J. J., Hennecke, M., Hutteman, R., Kandler, C., et al. (2014). What drives adult personality development? A comparison of theoretical perspectives and empirical evidence. *European Journal of Personality, 28*, 216–230.
- Steinberg, L., Albert, D., Cauffman, E., Banich, M., Graham, S., & Woolard, J. (2008). Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: Evidence for a dual systems model. *Developmental Psychology, 44*, 1764–1778.
- Tackett, J. L., Slobodskaya, H. R., Mar, R. A., Deal, J., Halverson, C. F., Baker, S. R., et al. (2012). The hierarchical structure of childhood personality in five countries: Continuity from early childhood to early adolescence. *Journal of Personality, 80*, 847–879.
- Terracciano, A., McCrae, R. R., Brant, L. J., & Costa, P. T. (2005). Hierarchical linear modeling analyses of the NEO-PI-R scale in the Baltimore Longitudinal Study of Aging. *Psychology and Aging, 20*, 493–506.
- Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.
- Van den Akker, A. L., Deković, M., Asscher, J., & Prinzle, 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*, 736–750.
- Van Lieshout, C. F. M., & Haselager, G. J. T. (1994). The Big Five personality factors in Q-sort descriptions of children and adolescents. In C. F. Halverson, G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 293–318). Hillsdale, NJ: Erlbaum.

APPENDIX

Extraversion (9 items)

18. Lets their peers know it when they are upset or angry. Doesn't hold back their own feelings when they feel upset or angry.
19. Is open and straightforward.
58. Openly shows the way they feel, whether it's good or bad.
84. Is a talkative person; talks a lot.
- 1R. Shows their thoughts and feelings in the way they look and act, but does not talk much about what they think and how they feel.
- 8R. Likes to keep their thoughts and feelings to themselves.
- 35R. Holds things in. Has a hard time expressing themselves; is a little bit uptight.

- 86R. Likes to be by themselves; enjoys doing things alone.
98R. Is shy; has a hard time getting to know people.

Agreeableness (22 items)

2. Is considerate and thoughtful of other people.
3. Is a warm person and responds with kindness to other people.
4. Gets along well with other people.
6. Is helpful and cooperates with other people.
9. Makes good and close friendships with other people.
14. Is eager to please.
15. Shows concern about what's right and what's wrong.
29. Is protective of others. Protects people who are close to them.
31. Is able to see how others feel; can put themselves in another person's place.
32. Gives, lends, and shares things.
62. Is obedient and does what they are told.
11R. Tries to blame other people for things they have done themselves.
13R. Tries to see what and how much they can get away with. Usually pushes limits and tries to stretch the rules.
20R. Tries to take advantage of other people.
22R. Tries to get others to do things by playing up to them. Acts charming in order to get their way.
55R. Worries about not getting their share of material things, food, or love. Seems afraid they won't get enough.
56R. Is jealous and envious; wants what other people have.
80R. Teases and picks on their peers.
85R. Is aggressive. (For example, picks fights or starts arguments.)
90R. Is stubborn.
93R. Is bossy and likes to dominate other people.
95R. Lets little problems get to them and is easily upset. It doesn't take much to get them irritated or mad.

Conscientiousness (10 items)

25. Thinks things out and reasons like a very mature person.
36. Finds ways to make things happen and get things done.
41. Is determined in what they do; does not give up easily.
47. Has high standards for themselves. Needs to do very well in the things they do.
59. Is neat and orderly in the way they dress and act.
66. Pays attention well and can concentrate on things.
67. Plans things ahead; thinks before they do something. "Looks before they leap."
76. Can be trusted; is reliable and dependable.
89. Is able to do many things well; is skillful.
99. Thinks about their actions and behavior; uses their head before doing or saying something.

Neuroticism (15 items)

23. Is nervous and fearful.
24. Worries about things for a long time.
33. Cries easily.
39. Freezes up when things are stressful, or else keeps doing the same thing over and over.
46. Tends to go to pieces under stress; gets rattled when things are tough.
48. Needs to have people say that they are doing well or ok. Is not very sure of themselves.
50. Tends to get sick when things go wrong or when there is a lot of stress. (For example, gets headaches, stomach aches, throws up.)
60. Gets nervous if they are not sure what's going to happen or when it's not clear what they are supposed to do.
72. Often feels guilty; is quick to blame themselves, even though they might not talk about it.
77. Feels unworthy; has a low opinion of themselves.
78. Has their feelings hurt easily if they are made fun of or criticized.
43R. Can bounce back and recover after a stressful or bad experience.
64R. Is calm and relaxed; easy-going.
82R. Speaks up and sticks up for themselves; goes after what they want.
88R. Is self-confident and sure of themselves; makes up their own mind.

Openness to Experience (6 items)

40. Is curious and exploring; likes to learn and experience new things.
68. Is a very smart person (even though formal tests and evaluations might not show this).
70. Daydreams; often gets lost in thought or a fantasy world.
74. Usually gets wrapped up in what they are doing.
96. Is creative in the way they look at things; the way they think, work, or play is very creative.
97. Likes to dream up fantasies; has a good imagination.

Activity (5 items)

26. Is physically active. Enjoys playing sports, running, and exercise.
28. Is energetic and full of life.
37. Likes to compete; is always testing and comparing themselves to other people.
51. Is well-coordinated. (For example, does well in sports.)
63. Is fast-paced; moves and reacts to things quickly.
Note. R indicates a reverse-keyed item.