Journal of Personality and Social Psychology
2006, Vol. 91, No. 3, 538–552

0022-3514/06/$12.00 DOI: 10.1037/0022-3514.91.3.538

Five Types of Personality Continuity in Childhood and Adolescence

Filip De Fruyt
Ghent University

Meike Bartels
Free University of Amsterdam

Karla G. Van Leeuwen, Barbara De Clercq, Mieke Decuyper, and Ivan Mervielde
Ghent University

This study examines 5 types of personality continuity—structural, mean-level, individual-level, differential, and ipsative—in a representative population ($N = 498$) and a twin and sibling sample ($N = 548$) of children and adolescents. Parents described their children on 2 successive occasions with a 36-month interval using the Hierarchical Personality Inventory for Children (I. Mervielde & F. De Fruyt, 1999). There was evidence for structural continuity in the 2 samples, and personality was shown to be largely differentially stable. A large percentage had a stable trait profile indicative of ipsative stability, and mean-level personality changes were generally small in magnitude. Continuity findings were explained mainly by genetic and nonshared environmental factors.

Keywords: personality continuity and assessment, five-factor model, childhood, adolescence, behavior genetics

In the past decades, there has been a wealth of studies on personality continuity that were recently summarized in two meta-analyses (MAs) of longitudinal data on differential (Roberts & DelVecchio, 2000) and mean-level (Roberts, Walton, & Viechtbauer, 2006) continuity. *Differential continuity* describes the degree to which the relative differences among individuals remains invariant across time, whereas *mean-level stability* refers to the extent to which personality scores change over time. Longitudinal designs are required to investigate differential stability, looking at trait correlations across time, whereas mean-level stability can be studied with the use of longitudinal data (Roberts et al., 2006). In addition, mean trait scores from cross-sectional age cohorts are useful for mean-level stability comparisons (McCrae et al., 1999, 2000).

A MA on differential continuity (Roberts & DelVecchio, 2000) showed that people acquire increasingly stable relative trait positions with age, with largely linear increases in stability until a plateau is reached at least after age 50. A MA on mean-level stability (Roberts et al., 2006) showed that people tend to increase, especially in their 20s to 40s, in social dominance (a facet of extraversion), conscientiousness, and emotional stability. People further demonstrate increases on social vitality (another facet of extraversion) and openness in adolescence, but decrease on both traits in old age. Absolute changes for agreeableness were observed only in old age. These two MAs convincingly illustrate that differential continuity and distinct mean-level change patterns can coexist, underscoring the independency of these two types of continuity (Block, 1971). Moreover, the two MAs combined help to advance a clear conceptual distinction between the two continuity types and are further useful to solve apparent inconsistencies among studies, summarizing the available data collected by the different researchers involved in the personality continuity debate. The two MAs raised new questions about additional forms of personality continuity and about the potential moderators and antecedents of stability.

Although the MA on rank-order continuity (Roberts & DelVecchio, 2000) provided stability estimates from 3 years to old age, the majority of studies in the MA for the youngest age groups relied on a limited set of temperament and Q-sort measures. Comprehensive and hierarchically organized age-specific personality measures might be more appropriate to study homo- and heterotypic developmental changes at young age (Caspi, 1998). The MA on mean-level changes (Roberts et al., 2006) started from age 10 onward, because there is a dearth of studies on this continuity type for younger ages. Today, the MAs allow for estimation of the amount of mean-level change and the stability correlation between any two ages from age 10 onward, but there is less evidence on younger age groups.

In addition to differential and mean-level continuity, other types of personality development have been examined, although less frequently (Caspi, 1998). Means on trait dimensions cannot be compared directly across measurement points when the covariance structure varies across time. *Structural continuity* refers to the invariance of the covariance structure across time and is a necessary requirement for the assessment of mean-level stability (Biesanz, West, & Kwok, 2003). In addition to the analysis of group means for different ages, one can also examine individual-level change. *Individual-level change* refers to the magnitude of increase or decrease exhibited by a person on any given trait. These changes may be masked in an analysis of mean-level con-

Filip De Fruyt, Karla G. Van Leeuwen, Barbara De Clercq, Mieke Decuyper, and Ivan Mervielde, Department of Developmental, Personality, and Social Psychology, Ghent University, Ghent, Belgium; Meike Bartels, Department of Biological Psychology, Free University of Amsterdam, Amsterdam, the Netherlands.

Correspondence concerning this article should be addressed to Filip De Fruyt, Department of Developmental, Personality, and Social Psychology, Ghent University, H. Dunantlaan 2, B-9000, Ghent, Belgium. E-mail: Filip.DeFruyt@ugent.be

tinuity, because equal numbers of individuals may increase or decrease on a trait, resulting in no change for the entire group. Finally, *ipsative stability* refers to the continuity of the configuration of traits within the individual and provides information on the stability of the patterning of traits within a person across time, hence facilitating a person-centered approach to personality development (Robins & Tracy, 2003). These additional types of personality continuity have been studied mainly from adolescence to adulthood (McCrae et al., 1999, 2000; Roberts & Del Vecchio, 2000; Robins, Fraley, Roberts, & Trzesniewski, 2001), although a few studies have addressed particular types of personality continuity in childhood (e.g., Van Lieshout & Haselager, 1994). Caspi, Roberts, and Shiner (2005) argued that there are relatively few studies that assess a comprehensive set of personality variables to track continuities and changes over time. No study has addressed all five types simultaneously across a substantial time interval in childhood and adolescence using a comprehensive and hierarchical five-factor model (FFM) personality measure.

The present study examines all five types of continuity in two different samples of schoolchildren and adolescents assessed at two time points spanning a 36-month interval. To assess whether the observed continuity and change patterns generalize across studies, we examined continuity types in a representative population sample and a genetic-informative sample of twins and siblings. The different nature of the samples further enabled us to investigate important additional questions. The representative sample allowed the description of patterns of continuity and change that are notable in the general population of children and adolescents. The twin and sibling sample allowed the examination of genetic and environmental influences on personality continuity and change, enabling a genetic–environmental decomposition of the personality trait variances cross-sectionally but also of the trait covariance across assessment points. McGue, Bacon, and Lykken (1993) conducted a similar study of young adulthood, demonstrating that the stable core of personality is strongly associated with genetic factors and that personality change largely reflects environmental factors. They reported that, on average, "over 80% of the variance of the stable component of the Time 2 phenotype was associated with genetic factors" but also that "a majority of the Time 2 personality variance is unrelated to variance expressed at Time 1" (McGue et al., 1993, p. 105). As far as we know, the present work is the first behavioral genetic study seeking to identify and characterize genetic and environmental influences on individual differences in stability and change in children and adolescents using a comprehensive FFM measure.

We examined personality continuity for a set of basic personality dimensions and facets using a lexically based measure assessing a broad range of personality traits relevant for childhood and adolescence. Adopting a comprehensive, age-appropriate inventory increases the likelihood of detecting changes and enhances the generalizability of findings. Although parents are the primary informants on children's and adolescents' personalities, the representative population sample also completed a self-report personality adjective measure so that we could examine shared method variance.

Temperament and Personality

Developmental and child psychologists traditionally describe stable and observable differences in young children by relying on temperamental constructs such as Negative Emotionality (Thomas & Chess, 1977) or Emotionality (Buss & Plomin, 1984), Sociability (Buss & Plomin, 1984) or Surgency (Rothbart & Derryberry, 1981), Task Persistence (Thomas & Chess, 1977) or Effortful Control (Rothbart & Derryberry, 1981), and Activity Level (Buss & Plomin, 1984; Goldsmith & Campos, 1982; Thomas & Chess, 1977). They assumed that temperamental differences are expressions of neurobiological mechanisms that have a strong genetic basis (Mervielde, De Clercq, De Fruyt, & Van Leeuwen, 2005). In contrast, personality traits are mainly used to chart stable latent differences in adults, presumed to be partly influenced by temperament and interaction with the environment.

The revival of trait psychology, and especially the preponderance of the FFM (Digman, 1990), challenged these viewpoints. First, personality psychologists tend to agree that five broad dimensions, that is, Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness, can be considered the basic dimensions underlying adult personality. Personality psychologists interested in the developmental antecedents of this FFM subsequently showed that these five are also useful to describe individual differences in childhood and adolescence (Digman, 1963; Digman & Inouye, 1986; John, Caspi, Robins, Moffit, & Stouthamer-Loeber, 1994; Kohnstamm, Halverson, Mervielde, & Havill, 1998; Lamb, Chuang, Wessels, Broberg, & Hwang, 2002). John et al. (1994) introduced the basic personality dimensions to developmental psychologists and demonstrated that the "Little Five" predict externalizing problem behavior in children. The qualifier "Little" refers to dimensions denoting trait differences in childhood, paralleling the label "Big" used to refer to dimensions characterizing adults.

Second, McCrae and colleagues (2000) argued against the artificial distinction between temperamental constructs and personality traits, because there are strong empirical and conceptual links between the domains of temperament and personality. The defining characteristics of temperament variables also apply to traits, including early observability, genetic basis, and pervasive impact on a range of behaviors (McCrae & Costa, 1997). Indeed, behavior genetic studies consistently document the strong genetic basis of FFM traits (Jang, McCrae, Angleitner, Riemann, & Livesley, 1998), with heritability estimates ranging from .40 to .60 depending on the trait and measure. McCrae and Costa (1996) conceptualized the FFM dimensions as basic tendencies shaping the interactions with the environment, resulting in characteristic (mal)adaptations such as interests, values, and attitudes but also problem behavior and psychopathology. Mervielde, De Clercq, et al. (2005) compared the major models of temperament with the FFM, illustrating the similarities and the conceptual overlap between temperament models and the FFM. Caspi et al. (2005) recently concluded that temperament and personality increasingly appear to be more alike than different.

Assessing Personality at Young Age

A comprehensive assessment of age-specific indicators of traits is crucial to studying personality at young age and especially its development. Different approaches have been adopted to assess FFM dimensions in children and adolescents. Often, FFM measures—initially developed for adults—are used to describe differences in younger age groups (see, e.g., studies using the Revised NEO Personality Inventory [NEO-PI–R] to assess adolescents' personality: De Clercq & De Fruyt, 2003; De Fruyt, Mervielde, Hoekstra, & Rolland, 2000; McCrae et al., 2002). Other studies adapted the phrasing of personality items for younger age groups (e.g., the junior version of Eysenck's Personality Questionnaire measure, EPQ–J; Eysenck, 1963; Eysenck, Makaremi, & Barrett, 1994). However, it can be argued that these adapted measures are probably not suitable for a fine-grained assessment of childhood and adolescent personality differences and especially not to assess developmental change (De Clercq, De Fruyt, & Van Leeuwen, 2004). Therefore, an alternative approach that is more sensitive to subtle personality differences at young age should be developed on the basis of the full range of personality differences observable prior to adulthood.

The lexical approach to personality description (De Raad, & Perugini, 2002) provides a convincing rationale for the development of a comprehensive child and adolescent personality taxonomy. Mervielde and De Fruyt (1999) adopted this approach to construct such a taxonomy for classification of a large pool of parental personality descriptions of Flemish children aged between 6 and 13 years (Kohnstamm et al., 1998). They subsequently developed the Hierarchical Personality Inventory for Children (HiPIC), representing the content of parental descriptions in short sentence items referring to concrete and observable behavior. Therefore, the HiPIC can be considered a lexically based measure of the active parental vocabulary, in contrast with the NEO-PI–R, in which the facets are not derived empirically but are selected after a careful search of the adult personality literature.

The HiPIC items span five broad domains, labeled as Extraversion, Benevolence, Conscientiousness, Emotional Instability or Neuroticism, and Imagination. Some domain labels differ from the lexical adult Big Five (Goldberg, 1993). The HiPIC dimensions Extraversion, Conscientiousness, and Emotional Instability refer to content that is similar to the adult Big Five counterparts and hence received the same label. The HiPIC Benevolence factor, however, refers to a broader set of traits than the adult Big Five or FFM Agreeableness factor because it includes traits linked to the "easy–difficult" child concept described in the temperament literature (Thomas, Chess, Birch, Hertzig, & Korn, 1963). The Benevolence factor refers to differences in the manageability of the child from the perspective of the parent-informant. The HiPIC Imagination domain comprises both Intellect and Openness to Experience items, blending the two alternative labels for the fifth factor emerging from adult adjective-based lexical studies (Goldberg, 1993) and the questionnaire-oriented FFM approach (Costa & McCrae, 1992). Given its specific focus and comprehensiveness, the HiPIC can be considered a more sensitive measure to assess personality change at young age (De Clercq et al., 2004).

Personality Development in Childhood and Adolescence

Different developmental theories conceive puberty as a significant stage for social and personality development (Grotevant, 1998), involving changing social interactions with parents and peers and increasing societal influences. The importance of adolescence as a key transitional phase is also acknowledged by Erikson (1950, 1968), who considered adolescence a second individuation stage in which change is likely. Similarly, behavioral theories (Robin & Foster, 1989) emphasize that during adolescence, learning processes and contingencies are embedded in novel social networks and environments, including changing peer groups. In addition, organismic theories such as Piaget's theory of cognitive development (Piaget, 1983) suggest that newly acquired cognitive structures influence the way children and adolescents interact with their environment. Finally, personality theories such as Cloninger's theory of character development (Cloninger, Svrakic, & Przybeck, 1993) discern qualitatively different life stages that an individual has to master before a more advanced developmental phase or level can be achieved. All these theories emphasize developmental discontinuities, but it remains to be established whether these discontinuities reflect changes in basic tendencies or whether they are restricted to changing characteristic adaptations (McCrae & Costa, 1996). The demonstration of different forms of stability across the FFM trait hierarchy in childhood and adolescence would underscore McCrae and Costa's Five-Factor Theory (McCrae & Costa, 1996) and hence requires that the previously reviewed theories of personality development account not only for trait change but also for trait stability.

Roberts and DelVecchio (2000) recently conducted an MA review of differential stability and examined whether and when stability peaks during the life course, challenging Costa and McCrae's claim (1994; McCrae & Costa, 1984, 1990) that personality is "set like plaster" after age 30. They analyzed 3.217 test–retest correlation coefficients from 152 longitudinal studies and demonstrated an increase in stability from .31 during childhood to .54 in young adulthood, mounting to .64 by age 30, with the highest stability of .74 observed between 50 and 70 years, when the time interval was held constant at 6.7 years. These data suggest that personality is less stable during the preadult years. The estimated population correlations controlling for the time interval of the longitudinal study are .31 for ages 0–2.9, .49 for ages 3–5.9, .43 for ages 6–11.9, and .43 for ages 12–17.9, respectively.

McCrae and colleagues (2002) examined both mean-level and differential changes during adolescence in self-descriptions using the NEO-PI–R, including an analysis of continuity at the level of the individual. They found that mean-level personality scores for Extraversion, Agreeableness, and Conscientiousness were stable between ages 12 and 18, but Neuroticism appeared to increase in girls and Openness to Experience increased in both boys and girls. A 4-year longitudinal study of intellectually gifted students showed a considerable degree of rank-order instability for gifted students across the two assessment points. Stability coefficients for boys across a 4-year interval ranged from .31 (Agreeableness) to .49 (Conscientiousness) and for girls from .30 (Neuroticism) to .63 (Conscientiousness). Individual-level continuity analyses relying on the reliable change index (RCI; Jacobson & Truax, 1991) indicated that about 60% of the sample did not change over the 4-year interval for each of the FFM dimensions. A similar analysis for older college-age individuals (Robins et al., 2001) showed, however, that almost 80% of the individuals were stable on each of the FFM dimensions, suggesting that differential stability is significantly lower during adolescence.

The MA of Roberts et al. (2006) on mean-level change during adolescence (10–18 years) showed a small but insignificant increase for Openness to Experience (d = .23; K = 13, N = 2,911), and significant increases for Social Dominance (d = .20; K = 5, N = 1,700) and Emotional Stability (d = 16; K = 23, N = 10,557). No significant mean-level changes were reported for Social Vitality, Agreeableness, and Conscientiousness during adolescence.

Block (1971) and Hart, Hofmann, Edelstein, and Keller (1997) used the California Q-set to examine within-person (ipsative) change, but to our knowledge no patterns of trait change in children and adolescents were examined with an inventory specifically designed to assess the FFM. Studies investigating stability of personality prototype classification using three personality profiles, that is, resilient, under-, and overcontrolled children, provided partial evidence of ipsative stability (De Fruyt, Mervielde, & Van Leeuwen, 2002). However, it was unclear whether the instability of prototype classification for some individuals should be attributed to real developmental change measurement error or the procedure to derive prototypes (De Fruyt et al., 2002). Structural continuity has not been studied explicitly in childhood and adolescence.

Method

Participants

Two longitudinal samples of children and young adolescents were available for the present analyses.[1]

Representative population sample. The first sample included children participating in a follow-up study investigating parenting, children's personalities, and problem behavior at two assessment occasions separated by a 36-month interval. Participating families including 1 child were recruited by random sampling of elementary and secondary schools. The sample was stratified by province (East and West Flanders), region (rural or urban), school type (public, private, or Catholic schools) and grade (3rd, 4th, 5th, and 6th year) for elementary schools. For secondary schools, subject sampling was based on province (East and West Flanders), type of curriculum (vocational, technical, and general education), and grade (1st and 2nd year). Eighty percent of the elementary schools and 60% of the secondary schools granted permission to contact parents for this study. In case of refusal, schools were replaced with other randomly selected schools. A letter addressed to the parents informed them about the goal and the procedures of the research project. The response rate for parents of primary schoolchildren was 41%, and for parents with children in secondary schools, 39%. At Time 2, 82% of the families continued collaboration, 12% refused, and 6% could not be reached. This sampling method resulted in a well-balanced sample regarding socioeconomic status, gender, and age (Van Leeuwen, 2004).

Families were visited at home by a trained psychology student at each assessment point. The mother, father, and child were instructed to independently complete a series of questionnaires. Parents described their own personality using the authorized Dutch translation of the NEO-PI–R (Costa & McCrae, 1992) and the personality of their child(ren) using the HiPIC (Mervielde & De Fruyt, 1999). Children also provided self-ratings of their personality using an adjective inventory at Time 2. Only participants who participated at both Time 1 and Time 2 were included in the study, resulting in a sample of 498 families, including 238 boys and 260 girls. There were no mean-level personality differences between the dropouts and those who continued participation, although both mothers and fathers who continued participation scored lower on Neuroticism, and fathers also scored higher, on average, on Extraversion, Openness to Experience, and Agreeableness. Indeed, continuing participation depended more on the personality of the parent than on that of the child. The mean age of the participants was 10.9 years (SD = 1.8 years; range 7–15) at Time 1, and 13.9 years (SD = 1.8 years; range 10–18) at Time 2.

Twin and sibling sample. The second sample included parents and children participating in a small-scale twin family study, providing HiPIC ratings of their twins and 1 sibling, if eligible, per family. HiPIC data on two assessment points with a 36-month interval were available for 548 children, all between the ages of 5 and 14 years, with a mean of 8.65 years (SD = 2.11 years). The sample included 271 boys and 277 girls.

Complex and differential interdependencies existed within the twin and sibling sample at the level of the units of observation, that is, within monozygotic (MZ) and dizygotic (DZ) twin pairs and between twins and siblings, but also at the level of the informants, that is, parents rating 2 (in the case of twins only) or 3 (twins + 1 sibling) children of the same family. The interdependencies at the level of the units of observation included a mixture of interchangeable (Griffin & Gonzalez, 1995) and noninterchangeable (Gonzalez & Griffin, 1999) cases, with same-sex twins to be considered interchangeable and different-sex or twin-sibling pairs as distinguishable. Such interdependencies might affect both the nature of the relationships observed across time and the level of significance (Gonzalez & Griffin, 2000). Because very different and complex analyses per type of personality continuity would have to be conducted to control for all possible dependencies that exist within such a dataset, a random sample of 1 child or adolescent twin per family was extracted to make the results directly comparable with the results obtained from the other samples involved in the present research, but also in previous MA work. The random twin sample included 208 twin children (boys, N = 106; girls, N = 102), whose ages were on average 8.50 years (SD = 1.92).

All twins and siblings for whom complete data records were available across the two assessment moments were used for the genetic analyses, taking into account their interdependencies. The twin sample included 35 MZ male pairs; 44 MZ female pairs; and respectively, 33 male, 35 female, and 56 mixed-sex DZ pairs of twins. The average age of the twins was 8.34 years (SD = 1.94), and the mean age for the siblings was 9.79 years (SD = 2.34).

Questionnaires

HiPIC. Parents rated both samples at the two assessment occasions on the HiPIC (Mervielde & De Fruyt, 1999). The HiPIC includes 144 items grouped into 18 facets that are hierarchically organized under the higher order factors. Parents were instructed to "describe the child by the way he or she has most often behaved over the last year" by indicating on a 5-point Likert-type scale the degree to which each statement was characteristic of the child to be assessed, with scale anchors labeled as *barely characteristic, slightly characteristic, more or less characteristic, characteristic*, and *highly characteristic*. All items have a similar grammatical format and are formulated in the third-person singular, avoiding negations in the item and excluding personality-descriptive adjectives.

Facet labels directly reflect the nature of the parental free descriptions and are in some cases indicative of opposite poles of the domain scale they are assigned to, requiring the computation of reversed scores (R) before facets can be aggregated into a domain score. The HiPIC structure outlines, as follows (a sample item is included between parentheses): Extraversion consists of Energy ("bubbles with life"), Expressiveness ("shows feelings"), Optimism ("laughs through life"), and Shyness (R: facet score to be reversed; "needs time to get used to peers"); Benevolence comprises the facets Altruism ("defends the weak"), Dominance (R; "acts the boss"), Egocentrism (R; "takes him/herself into consideration first"), Compliance ("sticks to arrangements") and Irritability (R; "is quick to take offense"); Conscientiousness includes Concentration ("works with sustained attention"), Perseverance ("keeps at it when the going gets tough"), Orderliness ("takes care of his/her possessions") and Achievement motivation ("wants to be among the best"); Emotional Instability measures Anxiety ("is afraid of failure") and Self-Confidence (R; "has confidence in own abilities"); and finally, Imagination is composed of the facets Creativity ("can use

[1] Both samples were already used for other research purposes. The representative population sample was already used to study the interaction between child personality and parental behavior as predictors of problem behavior (Van Leeuwen, Mervielde, et al., 2004), and the twin and sibling sample for a study on the personality type approach (De Fruyt et al., 2002). However, for both these purposes no data were reported on the continuity of personality.

everyday things in a new way"), Intellect ("grasps the meaning of things quickly"), and Curiosity ("asks many 'why' questions"). Benevolence is conceptually and empirically related to the adult FFM domain of Agreeableness, whereas Imagination is associated with the Openness to Experience domain (De Fruyt et al., 2000).

The HiPIC's robust factor structure and high internal consistencies of domains and facets have been documented in various studies with clinical and nonclinical samples (Van Hoecke, De Fruyt, De Clercq, Hoebeke, & Van de Walle, 2006; Van Leeuwen, De Fruyt, & Mervielde, 2004; Van Leeuwen, Mervielde, Braet, & Bosmans, 2004; Vollrath & Landolt, 2005). Domain scale reliabilities for the representative population and twin and sibling samples ranged from .76 (Extraversion, Time 1; representative population sample) to .89 (Conscientiousness, Time 1; twin and sibling sample), and for the facet scales, from .77 (Self-Confidence, Time 2; representative population sample) to .91 (Intellect, Time 1; twin and sibling sample).

Questionnaire Big Five. The population sample also provided self-ratings on the Questionnaire Big Five (QBF; Gerris et al., 1998; Goldberg, 1992), a Dutch shortened version of Goldberg's 100 adjectives, at the second measurement point. The QBF includes 30 adjectives, 6 per FFM dimension, presented with a 7-point Likert-type scale. Scholte, van Aken, and van Lieshout (1997) and Dubas, Gerris, Janssens, and Vermulst (2002) have demonstrated that the QBF provides valid estimates of an individual's standing on the FFM dimensions. Scale reliabilities were .77, .78, .67, .82, and .88 for Emotional Stability, Extraversion, Resourcefulness (the Openness scale of the QBF), Agreeableness, and Conscientiousness, respectively (Van Leeuwen, 2004).

Informants

In most cases, both parents provided HiPIC ratings of the children in the population and the twin and sibling sample. Given their high intercorrelations, ranging from .57 (Emotional Instability, representative population sample) to .77 (Conscientiousness, representative population sample), father and mother ratings were averaged in both samples to obtain more reliable scores. Self-ratings on the QBF at follow-up were also available for all children and adolescents of the representative sample, enabling an examination of shared method effects.

Data Analytic Approach

The two samples are age heterogeneous (at Time 1, age ranged from 7 to 15 years in the representative sample and from 5 to 14 years in the random selection of twins sample). The within-sample age range is thus much larger than the time interval of 36 months between the two assessment points. To account for this large age range, we combined the representative sample and the random selection of twins (N = 682) and then grouped in the following age groups at the first assessment: 6–7 years (N = 88), 8–9 years (N = 183), 10–11 years (N = 201), and 12–13 years (N = 210) to make the analyses more developmentally informative. The five continuity types are examined primarily within these different age groups.

Results

Structural Stability

The demonstration of invariance of the correlation matrix across the two measurement occasions was a conditio sine qua non before other forms of continuity could be examined. Table 1 reports the intercorrelations among the Time 1 and Time 2 domain scores averaged across mothers and fathers for the four age groups. Similarly to Robins et al. (2001), we examined structural continuity using structural equation modeling comparing the fit of two models. The correlations among the five factors were freely estimated at each measurement occasion in the first model, whereas in the second model the intercorrelations were constrained to be equivalent across assessments. A significant difference in the fit between these two models was considered indicative of structural change. The baseline model for the structural analysis at the Big Five factor level was a single-indicator latent variable model, with one latent variable associated with each of the 10 scores (five dimensions × two assessment occasions). This is a fully saturated model, with the variances of the latent variables fixed to 1 and the variances of the residuals fixed to 0. The correlations among the latent variables were freely estimated. In the second model, the correlations between all pairwise traits across the two assessment points were constrained to be equal. For example, the correlation between Extraversion and Emotional Instability at Time 1 was forced to equal the correlation between Extraversion and Emotional Instability at Time 2. Inspection of Table 2 shows that the intercorrelations among the five HiPIC domains were invariant across measurement occasions for all age groups: age group 6–7, $\chi^2\Delta(10, N = 88) = 15.69, p > .05$; age group 8–9, $\chi^2\Delta(10, N = 183) = 8.47, p > .05$; age group 10–11, $\chi^2\Delta(10, N = 201) = 9.52, p > .05$; and age group 12–13, $\chi^2\Delta(10, N = 210) = 11.45, p > .05$.

The previous analyses at the domain level could also be extended to examine structural invariance for the 18 facets across time. This very stringent test examining the invariance of 153 intercorrelations across time could not be done for the youngest age group because the number of parameters exceeds the number of participants. The 153 intercorrelations were not stable over time for age group 8–9, with $\chi^2\Delta(153, N = 183) = 207.27, p < .01$; or for age group 10–11, with $\chi^2\Delta(153, N = 201) = 198.34, p < .01$. Constraining the model for age group 12–13 showed no reduction in fit, with $\chi^2\Delta(153, N = 210) = 175.31, p > .05$, indicating that the 153 facet intercorrelations were invariant for the oldest age group across the 3-year interval.

Factor structures of varimax-rotated principal component analyses of the HiPIC facets could also be compared, calculating factor congruences across assessment occasions for each age group. Table 3 shows that the factor congruence coefficients were all higher than .90 except for the coefficients for Imagination from the age groups 8–9 and 10–11, which were .85 and .84, respectively.

Differential Continuity

Differential continuity coefficients across the 36-month interval (see Table 4), uncorrected for unreliability, were uniformly high for all domains across the four age groups, with values ranging from .61 (age group 6–7, Emotional Instability) to .86 (age group 6–7, Imagination), almost equaling test–retest reliabilities reported in the manual (Mervielde, De Fruyt, & De Clercq, 2005). For the first three age groups, coefficients for Emotional Instability were smaller than those for the other FFM domains, and the magnitude of the correlations decreased slightly with increasing age for Extraversion ($p < .01$ between the youngest and the oldest age groups). Very similar conclusions could be drawn for the HiPIC facets, with values ranging between .57 (age group 6–7, Anxiety) and .83 (age group 6–7, Altruism), suggesting that considerable differential stability is manifested across all hierarchical levels of the FFM.

Cross-time correlations computed separately for maternal and paternal ratings (not reported in Table 4) were about .05 to .10

Table 1
Intercorrelations Among the Five HiPIC Dimensions at Time 1 and Time 2 by Age Group

HiPIC dimension	1. Emotional instability	2. Extraversion	3. Imagination	4. Agreeableness	5. Conscientiousness
			Age 6–7 years		
1. Emotional Instability	—	−.33**	−.41***	−.05	.04
2. Extraversion	−.47***	—	.28**	.19	.00
3. Imagination	−.36***	.45***	—	.14	.42***
4. Agreeableness	.03	−.04	−.01	—	.42***
5. Conscientiousness	.02	−.09	.38***	.39***	—
			Age 8–9 years		
1. Emotional Instability	—	−.40***	−.27***	−.18*	−.04
2. Extraversion	−.40***	—	.48***	.03	−.04
3. Imagination	−.33***	.46***	—	.15*	.40***
4. Agreeableness	−.07	−.02	.01	—	.50***
5. Conscientiousness	−.03	−.02	.35***	.44***	—
			Age 10–11 years		
1. Emotional Instability	—	−.40***	−.37***	−.16*	−.20**
2. Extraversion	−.37***	—	.48***	.03	.04
3. Imagination	−.43***	.45***	—	.16*	.48***
4. Agreeableness	−.10	−.00	.07	—	.47***
5. Conscientiousness	−.22***	−.01	.45***	.42***	—
			Age 12–13 years		
1. Emotional Instability	—	−.44***	−.39***	−.19**	−.17*
2. Extraversion	−.43***	—	.52***	.13	.22***
3. Imagination	−.44***	.49***	—	.22**	.54***
4. Agreeableness	−.13	.06	.10	—	.49***
5. Conscientiousness	−.20**	.17*	.52***	.33***	—

Note. HiPIC = Hierarchical Personality Inventory for Children. Intercorrelations at Time 1 are reported below the diagonal, and intercorrelations at Time 2 are reported above the diagonal; 6–7 years: $N = 88$, 8–9 years: $N = 183$, 10–11 years: $N = 201$, and 12–13 years: $N = 210$.
* $p < .05$. ** $p < .01$. *** $p < .001$.

lower than the averaged parental ratings for domains and facets, suggesting that the magnitude of the stability coefficients does not primarily result from the increased reliability to be expected from averaging across parents. The continuity coefficients computed for single maternal ratings could be corrected for unreliability using 12-week test–retest reliabilities described in the manual (Mervielde, De Fruyt, & De Clercq, 2005) for maternal raters: .72 (Emotional Instability), .74 (Extraversion), .83 (Imagination), .78 (Benevolence), and .82 (Conscientiousness). Adopting these corrections for unreliability provided averaged stability coefficients for maternal domain ratings of .93, .91, .89, and .86 for the age groups 6–7, 8–9, 10–11, and 12–13, respectively. Correcting the coefficients obtained for the facets showed averaged corrected facet stability coefficients of .92, .90, .88, and .85 for the age groups 6–7, 8–9, 10–11, and 12–13, respectively.

Table 2
HiPIC Domain and Facet-Level Structural Continuity Analyses by Age Group Across a 36-Month-Interval

Age group	Chi-square	*df*	*p*	CFI
		Domain level		
6–7 years	15.69	10	.109	0.98
8–9 years	8.47	10	.583	1.00
10–11 years	9.52	10	.483	1.00
12–13 years	11.45	10	.324	1.00
		Facet level		
6–7 years[a]				
8–9 years	207.27	153	.00230	0.99
10–11 years	198.34	153	.00797	1.00
12–13 years	175.31	153	.10457	1.00

Note. HiPIC = Hierarchical Personality Inventory for Children; 6–7 years: $N = 88$, 8–9 years: $N = 183$, 10–11 years: $N = 201$, and 12–13 years: $N = 210$; CFI = comparative fit index.
[a] Total sample size is smaller than the number of parameters.

Table 3
HiPIC Factor Congruence Analyses Across a 36-Month Interval per Age Group

Age group	EI	E	I	B	C
6–7 years	.90	.95	.90	.98	.95
8–9 years	.97	.96	.85	.98	.91
10–11 years	.97	.95	.84	.99	.98
12–13 years	.99	.98	.97	.99	.98

Note. HiPIC = Hierarchical Personality Inventory for Children. 6–7 years: $N = 88$, 8–9 years: $N = 183$, 10–11 years: $N = 201$, and 12–13 years: $N = 210$; EI = Emotional Instability; E = Extraversion; I = Imagination; B = Benevolence; C = Conscientiousness.

Table 4
HiPIC Differential Continuity Analyses by Age Group Across a 36-Month Interval

HiPIC domain and facet	6–7 years	8–9 years	10–11 years	12–13 years
Emotional Instability	.61	.65	.63	.69
Anxiety	.57	.62	.60	.65
Self-confidence	.61	.63	.66	.66
Extraversion	.83	.78	.76	.66
Energy	.81	.71	.76	.63
Expressiveness	.74	.75	.69	.67
Optimism	.75	.64	.70	.59
Shyness	.77	.78	.68	.70
Imagination	.86	.69	.77	.69
Creativity	.82	.72	.73	.65
Intellect	.82	.70	.76	.74
Curiosity	.80	.64	.74	.62
Benevolence	.77	.71	.79	.75
Altruism	.83	.73	.71	.69
Dominance	.72	.61	.78	.70
Egocentrism	.69	.63	.71	.69
Compliance	.66	.69	.73	.66
Irritability	.75	.69	.76	.70
Conscientiousness	.76	.82	.77	.74
Concentration	.77	.80	.75	.73
Perseverance	.72	.74	.73	.67
Order	.68	.80	.73	.73
Achievement Striving	.72	.76	.72	.66

Note. HiPIC = Hierarchical Personality Inventory for Children; 6–7 years: $N = 88$, 8–9 years: $N = 183$, 10–11 years: $N = 201$, and 12–13 years: $N = 210$. All correlations significant at $p < .001$.

To evaluate effects of having the same informants (parents) and measures (HiPIC) across assessment occasions, we also examined differential stability using different measures *and* raters in the representative population sample. HiPIC-averaged parental ratings obtained at Time 1 were correlated with QBF self-ratings assessed 3 years later. Coefficients are provided in Table 5. To enable an evaluation of the effect of having different raters (self vs. parents) and measures, we also describe the correlations between HiPIC and QBF at follow-up. The follow-up coefficients primarily reflect the different informant and measures' perspectives, with coefficients, uncorrected for unreliability, varying between .21 (Benevolence–Agreeableness) and .47 (Conscientiousness), with all coefficients on the diagonal showing the largest magnitude. These coefficients are in line with other studies using self- and parental reports in children and adolescents (De Clercq, De Fruyt, Koot, & Benoit, 2004). The coefficients across the 36-month interval were in general about .10 lower than the same-time coefficients, with again all coefficients at the diagonal being largest.

Table 5
Cross-Method/Informant Correlations Across 36 Months (Representative Population Sample)

HiPIC domain	1. QBF–EmS	2. QBF–Ext	3. QBF–Res	4. QBF–Agr	5. QBF–Con
3-year interval					
1. HiPIC–EI	−.27***	−.22***	−.07	.02	.06
2. HiPIC–E	.07	.33***	.19***	.11*	−.08
3. HiPIC–I	.08	.09*	.26***	.05	.02
4. HiPIC–B	.10*	−.05	−.01	.12**	.11*
5. HiPIC–C	.05	−.01	.07	.08	.36***
Same-time assessment					
1. HiPIC–EI	−.38***	−.31***	−.12**	−.02	.07
2. HiPIC–E	.15***	.47***	.23***	.19***	−.01
3. HiPIC–I	.13**	.13**	.35***	.10*	.10
4. HiPIC–B	.13**	−.04	−.03	.21***	.24***
5. HiPIC–C	.03	−.05	.07	.17***	.47***

Note. HiPIC = Hierarchical Personality Inventory for Children; QBF = Questionnaire Big Fire; QBF–EmS = Emotional Stability; QBF–Ext = Extraversion; QBF–Res = Resourcefulness; QBF–Agr = Agreeableness; QBF–Con = Conscientiousness; HiPIC–EI = Emotional Instability; HiPIC–E = Extraversion; HiPIC–I = Imagination; HiPIC–B = Benevolence; HiPIC–C = Conscientiousness.
* $p < .05$. ** $p < .01$. *** $p < .001$.

The differences in magnitude between the different-measures and different-raters' correlations across time—relative to the same-measure parental correlations across the 36-month interval—are thus largely attributable to the different informants and measures, rather than testimony of poor differential continuity across this time lag.

McCrae (1994) showed by path analytic arguments that the stability of the true score can be easily estimated by dividing the predictive correlation by the concurrent correlation. The concurrent correlation between observers (self vs. parent) and/or different instruments (HiPIC vs. QBF) sets an upper limit to the agreement that can be expected, and the closer the cross-time correlation across observers is to this concurrent correlation, the more stable the trait (McCrae, 1994, p. 163). Adopting this estimation method to the present data shows true score validity estimates of .71 (.27/.38), .70 (.33/.47), .74 (.26/.35), .57 (.12/.21), and .77 (.36/.47) for Emotional Instability, Extraversion, Imagination, Benevolence, and Conscientiousness, respectively.

Mean-Level Continuity

We examined mean-level stability using repeated measures analysis of variance, including gender as a covariate (Costa, Terracciano, & McCrae, 2001; McCrae et al., 2002). The analyses of mean-level domain differences presented in Table 6 demonstrate strong parallels across the youngest age groups, with no mean-level differences reported for the age groups 6–7 and 8–9 for each of the Big Five. Small mean-level decreases in Emotional Instability were found for age groups 10–11, $F(1, 199) = 9.57$, $p < .01$, $\varepsilon^2 = .05$, and 12–13, $F(1, 208) = 9.86$, $p < .01$, $\varepsilon^2 = .05$; and decreases in Imagination, $F(1, 208) = 10.11$, $p < .01$, $\varepsilon^2 = .05$, and Conscientiousness, $F(1, 208) = 4.13$, $p < .05$; $\varepsilon^2 = .02$, were found for age group 12–13. No mean-level domain changes were found for Extraversion and Benevolence.

Mean-level stability analyses at the fact level (not reported in a table) showed a mean decrease for Dominance, $F(1, 86) = 5.27$, $p < .05$, $\varepsilon^2 = .06$, Optimism, $F(1, 86) = 6.42$, $p < .05$, $\varepsilon^2 = .07$, and Creativity, $F(1, 86) = 6.89$, $p < .01$, $\varepsilon^2 = .07$, for age group 6–7; a mean increase for Altruism, $F(1, 181) = 5.46$, $p < .05$, $\varepsilon^2 = .03$, for age group 8–9; and mean decreases for Energy, $F(1, 199) = 10.87$, $p < .001$, $\varepsilon^2 = .05$, Anxiety, $F(1, 199) = 12.10$, $p < .05$, $\varepsilon^2 = .06$, and Creativity, $F(1, 199) = 48.55$, $p < .001$, $\varepsilon^2 = .20$, for age group 10–11. The largest number of changes was found for age group 12–13, with mean decreases for facets across all domains: Irritability, $F(1, 208) = 11.52$, $p < .001$, $\varepsilon^2 = .05$; Achievement Striving, $F(1, 208) = 13.31$, $p < .001$, $\varepsilon^2 = .06$; Energy, $F(1, 208) = 4.84$, $p < .05$, $\varepsilon^2 = .03$; Expressiveness, $F(1, 208) = 7.62$, $p < .01$, $\varepsilon^2 = .04$; Anxiety, $F(1, 208) = 17.91$, $p < .001$, $\varepsilon^2 = .08$; Creativity, $F(1, 208) = 15.31$, $p < .001$, $\varepsilon^2 = .07$; and Curiosity, $F(1, 208) = 14.41$, $p < .001$, $\varepsilon^2 = .07$. In general, the magnitude of these differences was small, except for the decrease in Creativity for age group 10–11, and the majority of the observed differences were not significant after application of the Bonferroni correction for the number of statistical tests.

Individual-Level Continuity

We also examined whether mean-level continuity extended to the individual level, examining the number of individuals showing decreased, equal, or increased trait scores, using the RCI[2] (Christensen & Mendoza, 1986; Jacobson & Truax, 1991). Robins et al. (2001) and Roberts, Caspi, and Moffitt (2001) used this RCI to examine reliable change across the FFM traits. The RCI has been developed to assess the clinical significance of change after therapeutic intervention and is computed as $\mathrm{RCI} = X_2 - X_1 / S_{\mathrm{diff}}$, where X_1 represents a person's score at Time 1, X_2 represents that same person's score at Time 2, and S_{diff} is the standard error of difference between the two test scores. Computing RCIs for a person's Big Five profile enables researchers to determine how many individuals remain stable on their five-trait pattern across time but also provides information on the frequency of FFM individual-level change patterns. A child can be stable on four of the FFM dimensions but decline on Emotional Instability. This type of analysis is particularly useful to examine the kind and direction of individual changes but is also informative to describe the denser or more frequent trait-change configuration patterns in a large sample. Frequent individual trait-change configurations point to the kind of changes that can be expected during particular life stages, especially when replicable across samples of the same age range.

Difference scores per individual are compared with the distribution of change scores that would be expected from error of measurement alone and hence separates true change from change attributable to measurement error. Such analyses require independent test–retest reliability estimates. For the present analyses, HiPIC test–retest estimates for maternal[3] ratings across a 12-week interval reported in the manual (Mervielde, De Clercq, et al., 2005) were used, that is, .72 (Emotional Stability), .74 (Extraversion), .83 (Imagination), .78 (Benevolence), and .82 (Conscientiousness). Change scores exceeding a 95% confidence interval are assumed to represent true change (McCrae et al., 2002; Robins et al., 2001).

The majority of participants did not change FFM positions during the 36-month interval across the different age groups. Sixty-seven percent (age group 8–9 years) to 77% (age group 6–7 years) were ascribed similar FFM scores at the two assessment points. If children or adolescents changed on personality scores, change was usually restricted to one (about 20% across age groups) or two (about 5%–10%) FFM domains. None of the participants in the total sample ($N = 682$) changed on each of the FFM domains, and only 1 adolescent of age group 12–13 changed scores on four of the FFM domains. Some trait change patterns were denser across age groups, with a substantial number of individuals (more than 10%) from age groups 8–9 and 12–13 years exhibiting decreased scores on Imagination.

Ipsative Continuity

Ipsative continuity is usually examined with two methods. The first approach relies on Cronbach and Gleser's observation (1953)

[2] There has been some debate about the appropriateness of the RCI to account for regression to the mean effects when used to evaluate the effects of therapeutic interventions. However, this criticism applies less to the present application of the RCI, because no explicit assumptions are made regarding the direction of the expected change. Individuals can become, for example, more extraverted or more introverted.

[3] An anonymous reviewer pointed out that we do not have short-term test–retest correlations for HiPIC ratings averaged across parents, which is true. However, as we demonstrated, single parental ratings are only about .05 to .10 smaller across a 3-year interval than averaged ratings, so it can be expected that differences are even smaller across a shorter interval.

Table 6
HiPIC Domain Mean-Level Continuity Analyses by Age Group Across 36 Months

Domain and age group	Time 1	Time 2	*F*	*p*	ε^2
Emotional Instability					
6–7 years	22.41	22.79	.36	*ns*	
8–9 years	21.85	20.57	.76	*ns*	
10–11 years	21.39	20.32	9.57	.01	.05
12–13 years	21.69	20.11	9.86	.01	.05
Extraversion					
6–7 years	29.17	28.67	2.33	*ns*	
8–9 years	28.85	28.15	.22	*ns*	
10–11 years	28.70	27.63	1.29	*ns*	
12–13 years	27.82	26.89	2.14	*ns*	
Imagination					
6–7 years	30.38	29.25	.92	*ns*	
8–9 years	30.67	29.45	2.57	*ns*	
10–11 years	29.86	28.59	3.31	*ns*	
12–13 years	28.68	27.58	10.11	.01	.05
Benevolence					
6–7 years	26.56	26.94	.83	*ns*	
8–9 years	27.90	28.42	.03	*ns*	
10–11 years	28.00	28.44	1.31	*ns*	
12–13 years	27.90	27.92	1.88	*ns*	
Conscientiousness					
6–7 years	26.20	25.66	.01	*ns*	
8–9 years	26.59	26.28	1.13	*ns*	
10–11 years	26.24	26.02	.03	*ns*	
12–13 years	25.80	25.22	4.13	.05	.02

Note. 6–7 years: $N = 88$, 8–9 years: $N = 183$, 10–11 years: $N = 201$, and 12–13 years: $N = 210$; *F* according to Wilks's lambda; degrees of freedom for all analyses are (1, 86) for 6–7 years, (1, 181) for 10–11 years, and (1, 208) for 12–13 years; *ns* = nonsignificant. HiPIC = Hierarchical Personality Inventory for Children.

that individual profiles can vary in three major ways: elevation (the average level of scores), scatter (the variability of scores), and shape (the patterning of scores). Cronbach and Gleser (1953) developed three indices, D^2, D'^2, and D''^2, for quantifying these sources of variance. D^2 is sensitive to differences in elevation, scatter, and shape and quantifies the squared differences between Big Five traits at two assessment occasions. D'^2 is only sensitive to differences in scatter and shape and quantifies the squared differences between Big Five profiles after each profile has been centered around its mean. Finally, D''^2 only reflects differences in shape and quantifies the squared differences between profiles after each profile has been standardized (Cronbach & Gleser, 1953; Robins et al., 2001). Using these profile similarity indices, Robins and colleagues (2001) demonstrated that only 17% of their undergraduate sample showed significant changes in the shape of their profile across a 4-year interval. The majority of the changes in profiles were related to changes in elevation and/or scatter but not shape.

Parallel to the study by Robins et al. (2001, p. 626), probabilities were estimated by simulating trait scores on a sample of 50,000 individuals with identical levels of elevation, scatter, and shape in a person's profile at the two measurement points and examining corresponding distributions. Simulated trait scores were based on means, variances, and covariances estimated from the real data, and the test–retest reliability coefficients reported in the manual (Mervielde, De Clercq, et al., 2005) were used to estimate the error variance.

Across the four age groups, the D^2 values ranged from 35.70 to 45.68, the D'^2 values from 22.88 to 30.46, and the D''^2 from .20 to .22. To interpret these values, we simulated four samples of 50,000 participants, assuming no reliable changes in profiles across time: that is, with similar scores for elevation, scatter, and shape per age group. This simulation produced distributions for the D^2, D'^2, and D''^2 indices per age group, with the 95th percentiles for age group 6–7 years at 98.65, 84.46, and 1.31, respectively; for age group 8–9 years at 89.62, 76.16, and 1.19, respectively; for age group 10–11 years at 93.24, 79.11, and 1.20, respectively; and finally, for age group 12–13 years at 108.60, 92.58, and 1.39, respectively. Individuals with values beyond these 95th percentile values were considered to have significantly changed profiles. Respectively, 9.1%, 14.8%, 14.9%, and 16.7% of the children per age group had D^2 values beyond the simulated cutoffs; 5.7%, 10.4%, 11.4%, and 12.9% had D'^2 values beyond the simulated cutoffs; and 9.1%, 8.2%, 6.0%, and 9.0% had D''^2 values beyond the respective cutoffs, suggesting that children's and adolescents' profiles primarily reflected changes of elevation and scatter, but less so in terms of shape. Less than 10% of the individuals across all age groups exhibited changes in the shape of the profile.

A second but related approach to examining ipsative stability is to compute Q correlations, that is, within-person correlations across the HiPIC domains or facets at Time 1 and Time 2. Robins et al. (2001) found that Big Five profile correlations ranged from .95 to .97 during college years, with a mean of .61 ($SD = .39$) and a median of .76. Within-person correlations have to be evaluated against the distribution of within-person correlations that can be found in a sample with a similar mean and standard deviation, but in which profiles are randomly paired across assessment occasions. Robins et al. (2001) found an average value of .20 for a simulated data set.

Q correlations across the 18 HiPIC facets were computed per age group and compared with the distributions of Q correlations for simulations randomly pairing Time 1 with Time 2 assessments (i.e., data were from the same group of persons, but scores were paired randomly across time) per age group, having identical means and standard deviations at Time 1 and Time 2 as the real data. Across the four age groups, the median Q correlations ranged from .81 to .85, far above the median Q correlations in the simulated data (.35 to .43), suggesting stability of the within-individual facet-trait profile for a large number of children and adolescents.

Behavior Genetic Determinants of Continuity

Saturated model. Because of the small sample size, we combined MZ twins and DZ twins across gender to increase the power of the study. If nonsignificant effects of sex on variances and covariances are found, MZ male and MZ female twins can be combined in one group, and DZ male, DZ female, and DZ twins of opposite sex can be combined in one group. To allow sex differences in mean, we implemented sex as a covariate in the genetic model fitting. Tests were conducted with a saturated model in Mx (see http://www.psy.vu.nl/mxbib).

Genetic model fitting. Cross-sectional heritabilities and genetic and environmental influences on stability of the domains were estimated with the use of bivariate genetic analyses in Mx (Neale, 2003). A Cholesky decomposition, with sex and age as covariates (definition variables), was used. The Cholesky decomposition is descriptive and not driven by a specific developmental hypothesis. It decomposes a covariance matrix into genetic and nongenetic covariance matrices and is a first approach to obtaining genetic and environmental correlations across time in longitudinal datasets. Because of the small sample sizes, the analyses lacked power to test for the significance of the distinct variance components. However, significant effects of sex (girls showed higher ratings of Emotional Instability, Benevolence, and Conscientiousness at Time 1 and Time 2) and age (older children showed lower levels of Emotional Instability and Extraversion at Time 1 and Time 2) on the mean were found. Estimates for the full ACE model with the 95% confidence intervals are given.

No significant differences in variances and covariances for boys and girls were observed for any of the domains (p saturated model fitting is $> .05$ for all domains at the two time points), so for the remaining genetic analyses, MZ boys and girls were combined in one group, and DZ boys and girls (including twins of opposite sex) were combined in one group.

Standardized estimates (with 95% confidence intervals) of additive genetic, shared environmental, and nonshared environmental influences on all domains at Time 1 and Time 2 and on the covariance between Time 1 and 2 are given in Table 7. Heritability for the distinct domains varied from 7% for Emotional Instability at the second assessment occasion and 43% for Imagination at Time 1. Relatively large influences of nonshared environment were observed, partly representing measurement error. Shared environmental influences were small, except for Imagination, but it should be kept in mind that this variance component is hard to detect, and insufficient power gives rise to underestimation of the effects of shared environment. Stability of the personality domains between the two assessment points was accounted for by additive genetic, shared environmental, and nonshared environmental influences (see the boldfaced items in Table 7). The additive genetic standardized estimate for the trait covariance across time varied between .19 (Emotional Instability) and .51 (Benevolence). Estimates for the cross-time trait covariance for the shared environment were usually small, except for Imagination (.34), and estimates for the nonshared environment ranged from .31 (Imagination) to .68 (Emotional Instability).

Genetic and environmental correlations (i.e. overlap in influencing genes or environmental factors) between the first and the second assessment points are given in Table 8. Genetic correlations and shared environmental correlations were 1.0 or almost 1.0

Table 7
Standardized Estimates of Additive Genetic, Shared Environmental and Nonshared Environmental Influences on Variances and Covariances

HiPIC domains	A at Time 1	A at Time 2	C at Time 1	C at Time 2	E at Time 1	E at Time 2
Emotional Instability						
Time 1	.23 (.005–.43)[a]		.11 (.00–.23)[d]		.66 (.49–.87)[g]	
Time 2	**.19** (−.02–.46)[c]	.07 (.00–.23)[b]	**.13** (.00–.27)[e]	.06 (.00–.16)[f]	**.68** (.45–.90)[h]	.87 (.71–.97)[i]
Extraversion						
Time 1	.39 (.16–.56)		.06 (.00–.16)		.55 (.40–.76)	
Time 2	**.40** (.12–.62)	.28 (.05–.49)	**.10** (.00–.23)	.10 (.00–.23)	**.50** (.32–.76)	.61 (.44–.83)
Imagination						
Time 1	.43 (.28–.57)		.29 (.17–.40)		.28 (.20–.41)	
Time 2	**.35** (.18–.52)	.29 (.11–.45)	**.34** (.21–.46)	.30 (.20–.41)	**.31** (.20–.46)	.41 (.29–.57)
Benevolence						
Time 1	.38 (.17–.54)		.00 (.00–.05)		.61 (.46–.83)	
Time 2	**.51** (.24–.69)	.40 (.16–.57)	**.02** (−.02–.10)	.07 (.00–.16)	**.47** (.30–.74)	.53 (.38–.76)
Conscientiousness						
Time 1	.36 (.14–.53)		.00 (.00–.07)		.64 (.47–.86)	
Time 2	**.46** (.20–.65)	.38 (.15–.55)	**.00** (.00–.07)	.00 (.00–.06)	**.54** (.35–.80)	.62 (.45–.85)

Note. Boldfaced values indicate covariances, and 95% confidence intervals are in parentheses. HiPIC = Hierarchical Personality Inventory for Children; A = heritability; C = shared environment; E = nonshared environment.
[a] Heritability Time 1; [b] Heritability Time 2; [c] Influence of A on covariance between Time 1 and Time 2.
[d] Shared environment Time 1; [e] Shared environment Time 2; [f] Influence of C on covariance between Time 1 and Time 2.
[g] Nonshared environment Time 1; [h] Nonshared environment Time 2; [i] Influence of E on covariance between Time 1 and Time 2.

Table 8
Genetic and Environmental Correlations for the HiPIC Domains

HiPIC domain and time	A at Time 1	A at Time 2	C at Time 1	C at Time 2	E at Time 1	E at Time 2
Emotional instability						
Time 1	1		1		1	
Time 2	1.00	1	1.00	1	.57	1
Extraversion						
Time 1	1		1		1	
Time 2	.94	1	1.00	1	.67	1
Imagination						
Time 1	1		1		1	
Time 2	.82	1	.94	1	.74	1
Benevolence						
Time 1	1		1		1	
Time 2	1.00	1	1.00	1	.63	1
Conscientiousness						
Time 1	1		1		1	
Time 2	.98	1	.97	1	.67	1

Note. HiPIC = Hierarchical Personality Inventory; A = heritability; C = shared environment; E = nonshared environment.

for all HiPIC domains, suggesting one underlying set of genes and one underlying set of shared environmental influences for personality at both assessment occasions. Nonshared environmental correlations varied from .57 to .74, suggesting that large parts overlap but that there are also time-specific nonshared environmental influences, presumably measurement error.

Discussion

The present study examined five types of personality continuity in children and adolescents as rated by their parents on a lexically based measure specifically designed to assess personality at young age. The present work extends two recent MAs on differential (Roberts & DelVecchio, 2000) and mean-level (Roberts et al., 2006) personality continuity, assessing the two continuity types for younger age groups and examining three additional types of continuity for children and adolescents. Moreover, genetic and environmental influences on trait continuity were also estimated.

Structural Continuity

The SEM analyses showed that the covariance among HiPIC domains was clearly invariant across a 36-month interval for the different age groups, providing evidence for structural stability at the five-factor level from childhood to late adolescence. The results from these analyses at the domain level were also confirmed by the congruency analyses comparing factor structures for the different age groups across time points. All congruence coefficients were larger than .85, a value considered indicative of structural replicability (Haven & ten Berge, 1977), except for a minor deviation (.84) for Imagination in the 10–11 age group.

The most stringent test of structural personality continuity, examining the longitudinal invariance of the 153 facet intercorrelations, was positive only for the oldest age group of 12–13 years, suggesting that structural invariance applies to both hierarchical levels of the HiPIC, at least from 12–13 to 15–16 years. Structural invariance at the HiPIC facet level could not be demonstrated for the 8–9-year-olds and the 10–11-year-olds, although the chi-square values were small considering the large number of parameters, and the CFI indices were 1.0 or close to 1.0.

The present analyses clearly underscored that the positioning of the major dimensions of personality in childhood and adolescence is stable across a substantial time interval for different ages in childhood and adolescence. The demonstration of structural continuity for these age groups is a prerequisite before ipsative, mean-, and individual-level personality continuity data can be adequately examined and interpreted (Biesanz et al., 2003) and hence further legitimates the interpretation of MA findings reported for adolescence (Roberts et al., 2006).

Differential or Rank-Order Continuity

The rank ordering of individuals across the 3-year interval was demonstrated to be very stable, with coefficients usually beyond .70—uncorrected for unreliability—for the domains of Extraversion, Imagination, Benevolence, and Conscientiousness. Differential continuity for Emotional Instability was somewhat lower but still above .60. These findings largely generalized across the different age groups, except for a decline in rank-order continuity for Extraversion with increasing age. Analyses at the facet-level produced very similar results. Slightly lower correlation coefficients were observed for maternal ratings. These could be further corrected for unreliability showing correlations, averaged across FFM domains, ranging from .93 for the 6–7-year-olds to .86 for the 12–13-year-olds.

The results obtained here differ in three respects from previous reports on rank-order continuity. First, much higher correlations were found than those described in the MA on rank-order stability of Roberts and DelVecchio (2000) and the self-report results described by McCrae et al. (2002). The coefficients obtained in the present work double the correlations reported for this life stage in the MA of Roberts and DelVecchio (2000), estimated to be .43 for

12-year-olds, and they are also much higher than the median 4-year stability coefficient of .38 reported by McCrae et al. (2002) in a sample of gifted youth. The present coefficients obtained from parental ratings are similar to the correlations reported for adults, underscoring that FFM traits—when rated by parents—show already differential stability at young age.

Second, correlation coefficients were similar across age groups, except for a decrease of differential stability for Extraversion in the oldest group. These findings do not confirm the expectation (Roberts & DelVecchio, 2000) that differential continuity will increase with age. According to the present data, individuals have a stable positioning relative to each other in elementary and the first years of secondary school, except for Extraversion, at least when parental reports are considered.

A final difference is that in adulthood, self- and peer-reported rank-order continuity provide more similar coefficients (Costa & McCrae, 1988), whereas the findings from the present study—compared with those reported by Roberts and DelVecchio (2000) and McCrae et al. (2002)—suggest large differences in rank-order stability of adolescent self-reports versus parental reports. It could be argued that the higher rank-order coefficients for parental ratings are due to the tendency of parents to retain a lasting image of their child despite real changes. Indeed, involving the same observers and measures across assessment occasions makes it difficult to disentangle observer from developmental effects. However, the analysis correlating HiPIC parental assessments at Time 1 with QBF self-ratings at follow-up showed that the lower cross-time correlations were mainly attributable to the different informant perspectives (parent vs. self-rating) and the different inventories (HiPIC vs. QBF), rather than questioning differential stability. Relying on path-analytic arguments (McCrae, 1994), we see that true estimates for all HiPIC domains were above .70, except for a lower value of .57 for Benevolence in a design with different raters but also different measures. These values are certainly higher than those meta-analytically computed by Roberts and DelVecchio (2000) but remain considerably lower than similarly estimated true-score stabilities in adults as reported by Costa and McCrae (1988). The lower validities of the self-reports in other studies might be alternatively explained by a limited insight in personality in early adolescence.

Mean-Level Stability

No mean-level changes in childhood (age groups 6–7 and 8–9 years) were reported at the HiPIC domain level. In line with the findings of Roberts et al. (2006), mean-level decreases in Emotional Instability were reported from age 10 onwards, and no changes were reported for Benevolence. Contrary to the Roberts et al. (2006) findings, Conscientiousness and Imagination showed small decreases in age group 12–13, whereas Roberts et al. found no change for Conscientiousness and found increased scores, although insignificant, for Openness to Experience.

Roberts et al. (2006) distinguished two facets of Extraversion—Social Vitality and Social Dominance—and reported a significant increase for Social Dominance in adolescence. Social Vitality reflects traits such as sociability, positive affect, gregariousness, and energy level and is probably closely related to the HiPIC's Extraversion factor, including Energy, Shyness (R), Expressiveness, and Optimism, whereas their Social dominance category is probably best reflected by the HiPIC's Dominance facet of Benevolence. No mean-level changes for Extraversion were found in the present work, and mean-level Dominance scores slightly decreased for the youngest age group in childhood but did not show normative changes thereafter. The majority of differences at the facet level were reported for the late adolescents in our sample, although the magnitude of these differences was small, except the decrease in Curiosity reported for the 10–11-year-olds. Curiosity and also Creativity showed a slight decrease in the 12–13-year-olds. These findings are counter to those of Roberts et al., who found an increase in Openness during adolescence. An anonymous reviewer suggested that perhaps some of the HiPIC Curiosity and Creativity items are less applicable for these older age groups and are hence less frequently endorsed. However, inspection of the content of the Curiosity and Creativity items suggests that all items are applicable to a broader age range, including adolescence.

Individual-Level Changes

The analysis of the individual trait-change patterns demonstrated that two thirds (age group 8–9) to three quarters (age group 6–7) of all individuals did not show reliable change on any of the FFM dimensions underscoring individual-level continuity for a substantial number of individuals. If change occurred, it was usually restricted to one FFM domain. Only one trait-change pattern occurred more frequently; that is, children and adolescents showing decreased scores on Imagination, especially in the 8–9-year-old group and the 12–13-year-old group. These observations are counter to the findings of Roberts et al. (2006) and McCrae et al. (2002).

The implications of the individual-level findings for the personality research agenda are twofold. First, the present findings underscore the necessity to study the determinants of stability (Caspi et al., 2005), because the majority of children and adolescents are ascribed stable positions on traits. Secondly, determinants of change have to be investigated (Mroczek & Spiro, 2003), although the absence of dense trait-change patterns in the present samples suggests that the nature and direction of trait changes vary widely within the population, underscoring the necessity to examine very large groups. If the divergent pattern of individual trait changes would be replicated in such large samples, this would suggest that trait change is largely specific for the individual or even random, rather than normative.

Ipsative Stability

Although both individual-level and ipsative continuity indices focus on the individual, they are psychometrically very different. The computation of the RCI to inspect individual-level changes requires a comparison with the standard deviation of the population, contrary to the analysis of ipsative stability that relies exclusively on the individual's scores (De Fruyt, Van Leeuwen, Bagby, Rolland, & Rouillon, 2006). Ipsative stability was examined using two procedures applied to different hierarchical levels of the FFM. Q correlations were computed across the individual's facet scores at the two assessment points, whereas Cronbach and Gleser's (1953) D^2, D'^2, and D''^2 indices were used to examine differences among HiPIC domain-trait profiles across time.

The results from both types of analyses strongly converged. Substantially higher Q correlations within individuals across time were observed than for simulated ratings, assuming a random pairing of Time 1 and Time 2 scores. Using Cronbach and Gleser's (1953) D indices, we found that less than 10% of the sample exhibited changes in the shape of their domain profile, suggesting that if change occurs, it is mostly change of elevation or scatter. Similar ipsative analyses conducted across a 4-year interval in undergraduate students (Robins et al., 2001) showed that 17% demonstrated changes in the shape of their profile, suggesting that individual-level changes are observed more frequently in young adulthood than in childhood or adolescence.

Behavior genetic determinants of continuity. The twin and sibling sample, although small in size, provided a unique opportunity to decompose trait variances cross-sectionally as well as trait covariances across assessment points. In line with previous behavior genetic studies (Caspi et al., 2005; Loehlin, 1992), we estimated the additive genetic variance for the HiPIC domains to be around .30, varying between .07 (Emotional Instability, Time 2) and .43 (Imagination, Time 1). Smaller estimates were obtained for the shared environment, except for Imagination, and the remaining variance was explained by the nonshared environment, including measurement error. The small variance explained by the shared environment for the HiPIC domains (except for Imagination) has to be interpreted with caution, because the small sample size makes it difficult to detect shared environmental influences. These results are much in line with the patterns observed in other behavior genetic studies on individual differences in personality in adulthood (Jang et al., 1998).

Most interesting were the analyses on the genetic–environmental decomposition of the trait covariances (representing stability), showing very similar patterns for Extraversion, Benevolence, and Conscientiousness, with 40% (Extraversion) to 51% (Benevolence) of the covariance across time accounted for by genetic factors, 0% to 10% explained by the shared environment, and around 50% by the nonshared environment. The covariance of Emotional Instability was largely explained by the nonshared environment and to a moderate extent by genetic factors and 13% by the common environment. Finally, the three variance components explained an almost equal amount of covariance for Imagination. The magnitude of the genetic and shared environmental correlations (1.0 or near 1.0) suggests that the same set of genes and shared environmental influences determines the continuity observed across time, whereas the lower values for the nonshared environment suggest that different environmental factors, which are unique to an individual, operate across time.

Strengths and limitations. The present study has a number of strengths, including the use of a hierarchical and comprehensive personality measure specifically designed to assess traits at young age, as well as the availability of a population-based representative sample and a genetic informative sample.

However, this work also has a number of limitations that should be taken into account when interpreting the results. First, continuity was examined across a limited time lag with the use of only one inventory. It could be argued that more change is to be expected across longer time intervals, and the observed trait-change patterns—or the absence of change—could be specific for the HiPIC. For example, the decrease in Curiosity in adolescence might be particular for the HiPIC. It cannot be excluded that more changes will be observed across a longer time interval and for different personality measures. Second, the continuity types were examined with parents as primary informants on children's and adolescents' personality. Useful extensions of this research could include involving teachers as informants for young children and asking adolescents to provide self-ratings. Third, the genetic-informative sample is relatively small, hampering the statistical power of the analyses. The analyses that were conducted are informative, considering the dearth of genetic-informative studies on individual differences in personality development in childhood and adolescence, but larger samples would enable a series of more detailed and powerful analyses, including the exploration of the FFM lower level traits, and behavior genetic analyses of different types of personality continuity, for example, the trait pattern across time. Fourth, a potential threat for longitudinal research is that more stable individuals do not drop out, and hence personality continuity is more likely to be demonstrated than change. However, continuity is demonstrated across all traits not only for traits related to dropout in longitudinal research, such as Conscientiousness (De Fruyt & Mervielde, 1999). In addition, prolonged participation in studies with children depends more on parental decisions. In that case, parents who are more likely to continue participation describe their children as more consistent, which is unlikely to be the case. Finally, the present study included only two assessment occasions. Additional assessment points, preferably across a longer time span, would allow the use of latent-growth curve modeling and provide better ways to handle measurement error and offer opportunities to address new questions.

In conclusion, the evidence for different types of personality continuity supports and extends previous research demonstrating that the level of continuity in childhood and adolescence is higher than often expected (e.g., compared with young adulthood; Roberts et al., 2006). Caspi and colleagues (2005) argued in this respect that personality trait development is not a continuity-versus-change proposition, but that continuity and change coexist. The major challenge for developmental theories will be to account not only for trait changes but also for trait continuity.

References

Biesanz, J. C., West, S. G., & Kwok, O.-M. (2003). Personality over time: Methodological approaches to the study of short-term and long-term development and change. *Journal of Personality, 71,* 905–941.

Block, J. (1971). *Lives through time.* Berkeley, CA: Bancroft.

Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits.* Hillsdale, NJ: Erlbaum.

Caspi, A. (1998). Personality development across the life-course. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 311–388). Wiley: New York.

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 assessing change in a single subject: An alteration of the RC index. *Behaviour Therapy, 17,* 305–308.

Cloninger, C. R., Svrakic, D. M., & Przybeck, T. R. (1993). A psychobiological model of temperament and character. *Archives of General Psychiatry, 50,* 975–990.

Costa, P. T., Jr., & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology, 54,* 853–863.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory and Five-Factor Inventory Professional Manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1994). Set like plaster? Evidence for the stability of adult personality. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change*? (pp. 21–40). Washington, DC: American Psychological Association.

Costa, P. T., Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology, 81,* 322–331.

Cronbach, L. J., & Gleser, G. C. (1953). Assessing similarity between profiles. *Psychological Bulletin, 50,* 456–473.

De Clercq, B., & De Fruyt, F. (2003). Personality disorder symptomatology in adolescence: A Five-Factor Model perspective. *Journal of Personality Disorders, 17,* 269–292.

De Clercq, B., De Fruyt, F., Koot, H. M., & Benoit, Y. (2004). Quality of life, perceived competence and traits in children surviving cancer: A state-trait multiple-rater perspective. *Journal of Pediatric Psychology, 29,* 579–590.

De Clercq, B., De Fruyt, F., & Van Leeuwen, K. (2004). A "little five" lexically-based perspective on personality disorder symptoms in adolescence. *Journal of Personality Disorders, 18,* 479–499.

De Fruyt, F., & Mervielde, I. (1999). RIASEC types and Big Five traits as predictors of employment status and nature of employment. *Personnel Psychology, 52,* 701–727.

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., Mervielde, I., & Van Leeuwen, K. (2002). The consistency of personality type classification across samples and five-factor measures. *European Journal of Personality, 16,* 57–72.

De Fruyt, F., Van Leeuwen, K. G., 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.

De Raad, B., & Perugini, M. (2002). Big Five factor assessment: Introduction. In B. de Raad & M. Perugini (Eds.), *Big Five assessment* (pp. 1–26). Göttingen: Hogrefe & Huber.

Digman, J. M. (1963). Principal dimensions of child personality as inferred from teacher's judgements. *Child Development, 34,* 43–60.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41,* 417–440.

Digman, J. M., & Inouye, J. (1986). Further specification of the 5 robust factors of personality. *Journal of Personality and Social Psychology, 50,* 116–123.

Dubas, J. S., Gerris, J. R. M., Janssens, J. R. M., & Vermulst, A. A. (2002). Personality types of adolescents: Concurrent correlates, antecedents, and type × parenting interactions. *Journal of Adolescence, 25,* 79–92.

Erikson, E. (1950). *Childhood and society.* New York: Norton.

Erikson, E. H. (1968). *Identity: Youth and crisis.* New York: Norton.

Eysenck, S. B. G. (1963). *Junior Eysenck Personality Inventory.* San Diego, CA: Educational and Industrial Testing Service and Human Services.

Eysenck, S. B. G., Makaremi, A., & Barrett, P. T. (1994). A cross-cultural study of personality—Iranian and English children. *Personality and Individual Differences, 16,* 203–210.

Gerris, J. R. M., Houtmans, M. J. M., Kwaaitaal-Roosen, E. M. G., De Schipper, J. C., Vermulst, A. A., & Janssens, J. M. A. M. (1998). *Parents, adolescents and young adults in Dutch families: A longitudinal study.* Nijmegen, the Netherlands: Institute of Family Studies.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4,* 26–42.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48,* 26–34.

Goldsmith, H. H., & Campos, J. J. (1982). Toward a theory of infant temperament. In R. N. Emde & R. J. Harmon (Eds.), *The development of attachment and affiliative systems* (pp. 161–193). New York: Plenum.

Gonzalez, R., & Griffin, D. (1999). The correlational analysis of dyad-level data in the distinguishable case. *Personal Relationships, 6,* 449–469.

Gonzalez, R., & Griffin, D. (2000). On the statistics of interdependence: Treating dyadic data with respect. In W. Ickes & S. Duck (Eds.), *The social psychology of personal relationships* (pp. 181–213). London: Wiley & Sons.

Griffin, D., & Gonzalez, R. (1995). Correlational analysis of dyad-level data in the exchangeable case. *Psychological Bulletin, 118,* 430–439.

Grotevant, H. D. (1998). Adolescent development in family contexts. In W. Damon& N. Eisenberg (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 1097–1149). Wiley: New York.

Hart, D., Hofmann, V., Edelstein, W., & Keller, M. (1997). The relation of childhood personality types to adolescent behavior and development: A longitudinal study of Icelandic children. *Developmental Psychology, 32,* 195–205.

Haven, S., & ten Berge, J. M. F. (1977). *Tucker's coefficient of congruence as a measure of factorial invariance: An empirical study* (Tech. Rep. No. 290 EX).Groningen, the Netherlands: University of Groningen.

Jacobson, N. 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.

Jang, K. L., McCrae, R. R., Angleitner, A., Riemann, R., & Livesley, W. J. (1998). Heritability of facet-level traits in a cross-cultural twin sample: Support for a hierarchical model of personality. *Journal of Personality and Social Psychology, 74,* 1556–1565.

John, O. P., Caspi, A., Robins, R. W., Moffitt, T. E., & Stouthamer-Loeber, M. (1994). The little 5—Exploring the nomological network of the 5-factor model of personality in adolescent boys. *Child Development, 65,* 160–178.

Kohnstamm, G. A., Halverson, C. F., Mervielde, I., & Havill, V. L. (1998). *Parental descriptions of child personality. Developmental antecedents of the Big Five? The LEA series in personality and clinical psychology.* Mahwah, NJ: Erlbaum.

Lamb, M. E., Chuang, S. S., Wessels, H., Broberg, A. G., & Hwang, C. P. (2002). Emergence and construct validation of the Big Five in early childhood: A longitudinal analysis of their ontogeny in Sweden. *Child Development, 73,* 1517–1524.

Loehlin, C. J. (1992). *Genes and environment in personality development.* London, UK: Sage.

McCrae, R. R. (1994). The counterpoint of personality assessment: Self-reports and observer ratings. *Assessment, 2,* 159–172.

McCrae, R. R., & Costa, P. T., Jr. (1997). Personality trait structure as human universal. *American Psychologist, 52,* 509–516.

McCrae, R. R., & Costa, P. T., Jr. (1984). *Emerging lives, enduring dispositions: Personality in adulthood.* Boston: Little, Brown.

McCrae, R. R., & Costa, P. T., Jr. (1990). *Personality in adulthood.* New York: Guilford.

McCrae, R. R., & Costa, P. T., Jr. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In J. S. Wiggins (Eds.), *The five-factor model of personality: Theoretical perspectives* (pp. 51–87). New York: Guilford.

McCrae, R. R., Costa, P. T., Jr., de Lima, M. P., Simoes, A., Ostendorf, F., Angleitner, A., et al. (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., Ostendorf, F., Angleitner, A., Hrebickova, M., Avia, M. D., et al. (2000). Nature over nurture: Temperament, personality and lifespan development. *Journal of Personality and Social Psychology, 78,* 173–186.

McCrae, R. R., Costa, P. T., Jr., Terraciano, A., Parker, W. L., Mills, C. J., De Fruyt, F., & Mervielde, I. (2002). Personality trait development from age 12 to 18: Longitudinal, cross-sectional, and cross-cultural analyses. *Journal of Personality and Social Psychology, 83,* 1456–1468.

McGue, M., Bacon, S., & Lykken, D. T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental Psychology, 29,* 96–109.

Mervielde, I., De Clercq, B., De Fruyt, F., & Van Leeuwen, K. (2005). Temperament, personality and developmental psychopathology as childhood antecedents of personality disorders. *Journal of Personality Disorders, 19,* 171–201.

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

Mervielde, I., De Fruyt, F., & De Clercq, B. (2005). *Handleiding Hiërarchische Persoonlijkheidsvragenlijst voor Kinderen* [*Manual Hierarchical Personality Inventory for Children*]. Ghent, Belgium: Ghent University, Department of Developmental, Personality, and Social Psychology.

Mroczek, D. K., & Spiro, A., III. (2003). Modeling intraindividual change in personality traits: Findings from the Normative Aging Study. *Journal of Gerontology Psychology Science, 58B,* 153–165.

Neale, M. C. (2003). *Mx: Statistical modeling* (3rd ed). Richmond: University of Virginia.

Piaget, J. (1983). Piaget's theory. In P. H. Mussen (Ed.), *Handbook of child psychology: Vol. 1, History, theory, and methods* (pp. 103–128). New York: Wiley.

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., & 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.

Robin, A. L., & Foster, S. L. (1989). *Negotiating parent–adolescent conflict: A behavioral–family systems approach.* New York: Guilford Press.

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.

Robins, R. W., & Tracy, J. L. (2003). Setting an agenda for a person-centered approach to personality development. *Monographs of the Society for Research in Child Development, 68,* 110–122.

Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In M. E. Lamb & A. L. Brown (Eds.), *Advances in developmental psychology* (pp. 37–86). Hillsdale, NJ: Erlbaum.

Scholte, R. H. J., van Aken, M. A. G., & van Lieshout, C. F. M. (1997). Adolescent personality factors in self-ratings and peer nominations and their prediction of peer acceptance and peer rejection. *Journal of Personality Assessment, 69,* 534–554.

Thomas, A., Chess, L. A., Birch, H. G., Herzig, M. E., & Korn, S. (1963). *Behavioral individuality in early childhood.* New York: New York University Press.

Thomas, A., & Chess, S. (1977). *Temperament and development.* New York: Brunner/Mazel.

Van Hoecke, E., De Fruyt, F., De Clercq, B. J., Hoebeke, P., & Van de Walle, J. (2006). Internalising and externalising problem behavior in childhood enuresis: A Five-Factor Model perspective. *Journal of Pediatric Psychology, 31,* 460–468.

Van Leeuwen, K. (2004). *Parenting and personality as predictors of child and adolescent internalizing and externalizing problem behavior.* Unpublished doctoral dissertation, Ghent University, Ghent, Belgium.

Van Leeuwen, K., De Fruyt, F., & Mervielde, I. (2004). A longitudinal study of the utility of the resilient, overcontrolled and undercontrolled personality types as predictors of children's and adolescents' problem behavior. *International Journal of Behavioral Development, 28,* 210–220.

Van Leeuwen, K. G., Mervielde, I., Braet, C., & Bosmans, G. (2004). Child personality and parental behavior as moderators of problem behavior: Variable- and person-centered approaches. *Developmental Psychology, 40,* 1028–1046.

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, Jr., G. A. Kohnstamm, & R. P. Martin (Eds.). *The developing structure of temperament and personality from infancy to adulthood* (pp. 293–318). Hillsdale NJ: Erlbaum.

Vollrath, M., & Landolt, M. A. (2005). Personality predicts quality of life in pediatric patients with unintentional injuries: A one-year follow-up study. *Journal of Pediatric Psychology, 30,* 481–491.

Received May 12, 2005
Revision received April 4, 2006
Accepted April 6, 2006 ■