Published in final edited form as:

Eur J Pers. 2013 January 1; 27(1): 96–105. doi:10.1002/per.1886.

Observed Personality in Childhood: Psychometric and Behavioral Genetic Evidence of Two Broad Personality Factors

Zhe Wang,

Virginia Polytechnic Institute and State University

Nan Chen,

Virginia Polytechnic Institute and State University

Stephen A. Petrill, and

Ohio State University

Kirby Deater-Deckard

Virginia Polytechnic Institute and State University

Abstract

We examined broad dimensions of children's personalities (total $n = 1056$; age = 3.5 to 12 years) based on observers' perceptions following a few hours of structured interaction. Siblings' behaviors during a two-hour cognitive assessment in the home were rated separately by two different observers. Exploratory and confirmatory factor analyses clearly revealed a two-factor solution in three different samples. There was correspondence between parent-rated temperament and the observer-rated factors. Cross-sectional analyses indicated lower Plasticity among older children and higher Stability among older children. Sex differences were negligible. Plasticity and Stability were correlated in the .2 to .3 range. Most of the sibling similarity in the Plasticity was due to additive genetic influences, whereas most sibling similarity in Stability was attributable to shared environmental influences. The findings implicate a biometric factor structure to childhood personality that fits well with emerging bio-social theories of personality development.

Keywords

broad personality factor; observation; childhood; psychometric; behavioral genetics

How many dimensions are necessary to describe individual differences in personality, and when does this structure emerge in development? For many years, the "Big Five" (i.e., Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) has been considered the most supported model (Costa & McCrae, 1992; Goldberg, 1993). However, there continues to be discussion of whether the Big Five is the most robust and broadest structure in the personality hierarchy (DeYoung, Peterson, & Higgins, 2002; Digman, 1997;

Correspondence concerning this article should be addressed to Kirby Deater-Deckard, Department of Psychology, 109 Williams Hall (0436), Blacksburg, VA, 24061. kirbydd@vt.edu.

Zhe Wang, Department of Psychology, Virginia Polytechnic Institute and State University; Nan Chen, Department of Psychology, Virginia Polytechnic Institute and State University; Stephen A. Petrill, Department of Human Development and Family Science, Ohio State University; Kirby Deater-Deckard, Department of Psychology, Virginia Polytechnic Institute and State University.

Musek, 2007; Saucier, 2003). In the current study, we examined the early emerging personality structure in a group of children ranging from 3.5 to 12 years of age using observers' perceptions following a few hours of structured interaction.

Beyond the Big Five

Although the Big Five personality structure has been considered the most supported model in explaining individual differences in personality, evidence supporting fewer and broader factors has been emerging in recent years. In particular, two higher order factors have been extracted and replicated across multiple studies (DeYoung, 2006; Digman, 1997). The first factor was labeled as "Stability" (Agreeableness, Conscientiousness, and Neuroticism), and represents the tendency to self-regulate to maintain a stable organization of psychosocial function. The second factor was labeled as "Plasticity" (Openness and Extraversion), and represents the willingness to take risk and the flexibility to incorporate novel information into personal growth.

Although these higher-order factors have been identified psychometrically, their theoretical and empirical meanings are only beginning to be understood (McCrae & Costa, 2008). The center of the debate lies in whether these higher-order factors are substantive constructs that reflect bio-social substrates or method artifacts. First, some researchers have argued that correlations among the lower-order personality factors are due to individual instrument bias (Costa & McCrae, 1992) or individual rater bias (Biesanz & West, 2004). However, more recent studies have successfully extracted higher-order factors using multiple instruments (Rushton et al, 2009) or multiple-informant ratings (DeYoung, 2006), suggesting that these higher-order factors are not a reflection of individual instrument or single informant bias.

The second argument against higher-order personality factors is that correlations among the Big Five are due to substantial secondary loadings at lower-level scales which represent blends of two or more Big Five factors. (Ashton, Lee, Goldberg, & De Vries, 2009). However, depending on the quantity of the lower-level scales with substantial secondary loadings and how they cluster, these lower-level scales may be telling us more than that they are simply poor measures. If most lower-level scales have substantial secondary loadings in a way that combines Extraversion with Openness, and combines Agreeableness, Conscientiousness, and Neuroticism, these lower-level scales may capture the common personality characteristics that cannot be described by any single lower-level factor. If so, this would suggest that broader personality factors exist that represent more general traits than the Big Five.

Third, some researchers have argued that there may still be biases like those proposed by Paulhus and John (1998) and McCrae and Costa (1999), namely moralistic bias and egoistic bias. Moralistic bias is an individual's overestimation of his/her own dutifulness and cooperativeness, and resembles the Stability factor. Egoistic bias, which resembles the Plasticity factor, is when individuals overstate their own social and intellectual status. However, instead of explaining these factors as nothing more than method artifacts, an alternative is that such a particular organization of "person-perception bias" reflects

common ways in which person-perception schema is organized in our memory according to which we perceive and understand other people and ourselves.

Supporting evidence of the above argument can be found in studies using a lexical approach. By examining the most frequently used personality-descriptors in a language, the result reflects generally how people perceive and understand others in that particular culture. Results from several lexical studies have suggested that the broad two-factor solutions can be consistently replicated across several languages (Goldberg & Somer, 2000; Saucier, 2003; Saucier, Georgiades, Tsaousis, & Goldberg, 2005). Similar to the Stability factor, the first factor was named Morality and includes personality descriptors such as industrious, considerable, and bad-tempered. The second factor, which resembles the Plasticity factor, was named Dynamism and includes adjectives such as energetic, expressive, clever and talented (Saucier et al., 2005).

The Functional Role of the Metatraits

One of the most important reasons why we care a great deal about personality is that stable individual differences in these personality traits help us predict other people's behaviors and make our own decisions concerning those people more precisely. Our knowledge about another person's personality may be general or specific, and the usefulness of a particular type of knowledge often depends on our closeness to the particular person in question. For example, accurate "if-then" knowledge about a close friend's personality characteristics could reduce feelings of conflict and increase feelings of depth in the relationship with that friend (Friesen & Kammrath, 2011).

However, there are also many other situations in life when a detailed personality knowledge of another person is unnecessary and inefficient. In those situations, a quick decision is usually needed based on what one has learned about the target person in a relatively short period of time. One example is when people make judgments of personality of politicians (Caprara, Barbaranelli, & Zimbardo, 1997). Using personality adjectives, it has been found that although people rated themselves and their familiar basketball players using all the Big Five factors, their ratings on politician's personality were restricted to two broad factors; one factor was a combination of agreeableness, conscientiousness, and neuroticism (i.e., stability), and the other factor was a combination of extraversion and openness (i.e., plasticity; Caprara et al., 1997). This example potentially suggests that when people have to make an important decision based on a large amount of information with time constraints, they are very likely to adopt a cognitively efficient strategy via a two-factor person-perception schema (i.e., Stability and Plasticity). Following this reasoning, one would hypothesize that when using direct behavioral observation during which time an observer only sees a target individual for a limited period of time, the ratings should capture the broader personality traits of the target person. The first aim of the current study was to test this hypothesis.

When do Metatraits Emerge?

One other limitation of the previous studies examining the two broad personality factors is that almost all studies have examined adult samples. Therefore, our second aim in the

current study was to address this gap to examine whether metatraits were evident in childhood to early adolescence.

In infancy and childhood, studies on individual differences in stable traits have focused on temperament, which refers to biologically influenced individual differences in reactivity and self-regulation that are observable from birth onwards and emerge as stable dispositions that operate across time and situations (Rothbart & Bates, 1998). Generally, researchers believe that there is continuity between child temperament and adult personality traits. Some researchers believe that temperament characteristics are the initial state from which adult personality develops (Rothbart, 2007), and that adult personality traits represent the social and cognitive elaborations of individual differences in early temperament (Caspi, 2000). There are also other researchers who have challenged the conceptual distinctions between temperament and personality traits. They have argued that the distinction between personality and temperament is obscure and that the separation between personality studies in adulthood and temperament studies in infancy and childhood is artificial (McCrae et al., 2000). One important challenge in linking infant and child temperament to adult personality lies in that the scope of observable individual differences in behaviors may change with maturation (Caspi, Roberts, & Shiner, 2005). For example, observable individual differences in more complex emotions such as guilt and pride do not emerge until about age 3 (Eisenberg, 2000). As a result of a relatively limited set of observable traits in early years, young children may exhibit fewer trait factors (Rothbart, Ahadi, Hershey, & Fisher, 2001). Despite the challenges, a number of studies have demonstrated that the trait factors that are observable in children are very similar to those seen in adults (Barbaranelli, Caprara, Rabasca, & Pastorelli, 2003; De Fruyt, Mervielde, Hoekstra, & Rolland, 2000; Shiner & Caspi, 2003), suggesting a continuity in personality development (Caspi et al., 2005).

Using parent- and self-reports, Slobodskaya (2011) has successfully extracted the two broad personality factors (i.e., Stability and Plasticity) in 3- to 17-year-old children and adolescents in Russia, providing the first evidence that the two broad personality factors concerning socialization and personal growth are as evident by childhood as they are in adulthood. In the current study, we attempted to extend this literature by examining the evidence for the early emergence of the two broad personality factors through childhood and early adolescence using stranger observer's ratings.

Are Broad Personality Factors Heritable?

To explain the substantive meanings of the broad personality factors, theories have been proposed regarding the role of underlying biological mechanisms. One proposed model links the two broad personality factors with the functioning of serotonergic and dopaminergic neurotransmitters systems (DeYoung et al., 2002). In terms of underlying genetic influences, there has been only one behavioral genetic study of the two broad personality factors (Jang, et al., 2006). Using an ethnically diverse sample of adolescents and adults, the researchers found moderate to substantial heritability ($h^2 = .36 - .72$) and non-shared environmental variance ($e^2 = .28 - .64$) for both factors; shared environmental variance was negligible. Due to the scarcity of behavioral genetic studies of these two broad personality factors—

especially in childhood and early adolescence—the third aim of the current study was to address this gap.

In summary, the aims of the current study were to address the following three questions:

1. Are the two broad personality factors observable by a stranger observer after a few hours of interaction?
2. Are the two broad personality factors already evident in childhood and early adolescence, and are there any mean level differences in the two broad factors across different ages and sexes?
3. What are the genetic and environmental influences on the two broad personality factors?

Methods

Samples

The current study utilized existing archived data from three behavioral genetic studies. The TRACKS twin study (Deater-Deckard, 2000) included 118 same-sex twin pairs who were 3.5 years old (59% female; 52% monozygotic or MZ). Nearly all were Caucasian English. Fifty to fifty-nine percent of parents had the British equivalent of a high school diploma or less education, 27–31% had college degrees, and 10–13% had postgraduate degrees. The WRRP (Western Reserve Reading Project; Petrill, Deater-Deckard, Thompson, DeThorne, & Schatschneider, 2006) twin study included 296 same-sex (59 % female) twin pairs (43% MZ) who were 6.09 years old on average ($SD = .69$, range = 4.33 to 7.92). Ninety percent of the sample was Caucasian, 5% African-American, and 2% Asian. Twenty-three to twenty-five percent of the parents had post-graduate education, 30–31% bachelor's degree, 22–27% some college, and 11–16% high school or less. The N2CAP (Northeast-Northwest Collaborative Adoption Projects; Deater-Deckard & Petrill, 2004) included 114 genetically unrelated sibling pairs who were selected from the larger dataset so that only those pairs in which both siblings were under 12 years of age (57% female) were included. Sibling 1 was 9.07 years old on average ($SD = 1.83$), and sibling 2 was 6.38 years old on average ($SD = 1.73$). Thirty-six percent of sibling 1 was Asian, 43% Caucasian, 4% African-American, and 8% Latino. Sixty percent of sibling 2 was Asian, 27% Caucasian, 5% African-American, and 6% Latino. Forty-six to sixty-four percent of the parents had post-graduate education, 20–22% had bachelor's degree, 12–20% had some college, and 4–5% had high school or less.

Procedures and Measures

In each study, a two to three hour long cognitive assessment was conducted during a home visit. At the end of the visit, each tester independently rated the behavior of the child she or he had just assessed using Bayley's (1969) Behavior Record (BBR) which includes 26 items rated on a 5-point Likert-type scale. Each child was rated only by one observer, which did not allow for estimation of inter-rater reliability.

The BBR was chosen as our observation instrument to measure child behaviors because it captures a variety of aspects of affect and behavior that pertain to a wide range of personality facets (e.g., positive and negative affect, interest, exploration, attention/persistence, social engagement). Detailed item descriptions were provided in the result section. The BBR has been applied by many previous studies to measure infant and child personality and temperament (Freedman & Keller, 1963; Matheny, Dolan, & Wilson, 1976). For example, Freedman & Keller (1963) have examined the development of personality and motor ability using the BBR. They assessed infant personality and motor ability during various tests on a monthly base during the 1st year of life and found genetic influence on both personality and motor ability in infant twins. Other studies have used the BBR as a criterion measure for development of new measures (Caspi, 2000; Goldsmith & Gottesman, 1981). Both of these studies have applied a measure (i.e., the longitudinal collaborative perinatal project (CPP)) that is similar in scope to the behavior ratings contained in the BBR. Similar to the use of the BBR, behavioral observation using the CPP was either performed in a testing session involving cognitive and motor tasks or in a free play session. Using a person-centered approach, Caspi (2000) has found that temperament assessed in 3-year-old children using the CPP predicted the development of personality and psychopathology in adolescence and adulthood. In the Goldsmith & Gottesman (1981) study, two, three, and four temperament factors were obtained using the CPP when the participants were 8 months, 4 years, and 7 years old, respectively. However, careful examination of the factor structures at 4 and 7 years reveals that about half of the items had double-loadings across two or more factors.

For examining the construct validity of observers' perceptions, we examined its correspondence with the widely-used temperament instruments as the criterion. We chose parents' ratings of child temperament because they were the most commonly used in childhood through early adolescence. Mother and father ratings were moderately correlated so were averaged. In the WRRP, parents completed the Children Behavior Questionnaire Short Form (CBQ-SF, Putnam & Rothbart, 2006). The CBQ measures three broad dimensions of child temperament (i.e. effortful control, negative affectivity, and surgency) on a 7-point Likert scale. Coefficient alpha for these scales ranged from .60 to .87. In the TRACKS and N2CAP studies, parents completed the Colorado Child Temperament Inventory (CCTI, Rowe & Plomin, 1977). Coefficient alpha could not be computed from the TRACKS and N2CAP archived datasets but previous studies with the CCTI have reported good internal consistency (Cronbach's alpha coefficient of .73 to .88; Rowe & Plomin, 1977). In terms of face validity, the CCTI persistence scale is similar to the CBQ effortful control scale, and the CCTI emotionality scale is similar to the CBQ negative affectivity scale. A computed composite of the CCTI sociability, activity level, and shyness (reversed) scales is similar to the CBQ surgency scale. The CCTI surgency composite was internally consistent in the TRACKS dataset (first principal component explained 52% of the variance, loadings > .62) and the N2CAP dataset (first component explained 64% of the variance, loadings > .73).

Results

Factor Analysis and Score Construct Validity

Factor Analysis—To identify an initial structure as the basis for subsequent confirmatory factor analysis, preliminary exploratory factor analysis was conducted using the N2CAP dataset, with separate analyses conducted for sibling 1 and sibling 2. The last 9 items on the BBR were excluded from analysis because they measured extremely rare motor phenomena (e.g., tremors, poor muscle tone) and had virtually no variance. Using principal axis factoring with oblique rotation (direct oblimin), a two-factor solution was clearly suggested from visual inspection of the scree plots. Parallel analysis was also performed to determine the number of factors (Zwick & Velicer, 1986). According to this method, eigenvalues derived from the actual data are compared to the eigenvalues derived from the random datasets that parallel the actual dataset with regard to the number of cases and variables. Factor is retained if the eigenvalue from the actual dataset is greater than the eigenvalue from the random datasets (O'Connor, 2000). Eigenvalues from the actual and random datasets are shown in Table 1, and clearly suggested a two-factor solution. Items that had double loadings were removed to obtain a simple structure (loadings > .30 on only one factor). The items that were removed included 'g' (interest in test materials and stimuli), 'h' (initiative with tasks), 'l' (enthusiasm toward tasks), and 'm' (fearfulness). The two factors accounted for 51% (sibling 1) and 48% (sibling 2) of the variance in the items. The inter-factor correlation was .25 (sibling 1) and .33 (sibling 2). Based on the apparent semantic correspondence between the items marking each factor and traits that have been associated with Stability and Plasticity, we named the first factor "Stability" ('b' [low negative affect], 'c' [soothability], 'd' [low hypersensitivity], 'f' [adaptation to change], 'j' [attention], 'k' [persistence], 'n' [low frustration], 'o' [oriented to observer], and 'q' [cooperation]). The second factor was named "Plasticity" ('a' [positive affect], 'e' [energy], 'i' [exploration], and 'p' [social engagement]).

Using Amos 18.0 (SPSS Inc., Chicago, IL, USA), confirmatory factor analysis (CFA) in structural equation modeling (SEM) was used to test for a two factor structure (Plasticity and Stability) in the other two datasets (TRACKS and WRRP). Because each dataset had a relatively small sample size, we combined the two twin datasets for the CFA (Curran & Hussong, 2009). Before performing the CFA, we first examined the comparability in means and standard deviations of the 13 BBR items across the two samples using multivariate analyses. The means and standard deviations of the 13 items are shown in Table 2. Multivariate analyses suggested that there were overall significant differences in means and standard deviations between the WRRP and TRACKS (overall test of mean difference: $F_{(13, 825)} = 37.67, p < .001$; Overall test of variance difference: $\chi^2 = 177.10, df = 13, p < .001$; detailed results available upon request). Because CFA is based on variance-covariance matrix patterns, mean difference across samples would not affect factor structures. To control for between-sample sources of variation, we included a dummy coded variable in the CFA model (Figure 1) to denote study membership of each individual as suggested by Curran & Hussong (2009) concerning the simultaneous analysis of multiple datasets. Sibling 1 and sibling 2 were analyzed together as interchangeable dyads using the method suggested by Olsen and Kenny (2006). The CFA model is shown in Figure 1. In this model, item "q"

and item “a” were fixed at 1 to set the scale of the matrix. Stability and Plasticity factors were correlated with each other. Correlations between residual variances were allowed within each factor only when the following two conditions were met: (1) modification indices suggested that the addition of these correlations would significantly improve model fit; and (2) these correlations made theoretical sense. As a result, five correlations between residual variances in items were added as shown in Figure 1. Residual variances in negative affect and soothability and residual variances in frustration and soothability were allowed to correlate because they all represented aspects of negative affectivity. Residual variances in attention and persistence were correlated as they both indicated attentive behaviors. Residual variances in orientation to observer and cooperation were allowed to correlate because they both assessed agreeableness. Lastly, correlation between residual variances in hypersensitivity and adaptation to change was added because both items measured individual’s adaptability to novelty. Factor loadings and model fit indices are shown in Table 3. All loadings were statistically significant ($p < .001$). Model fit was adequate (RMSEA $< .08$, TLI $> .90$; Akaike, 1974; Steiger, 1990; Tucker & Lewis, 1973), suggesting an acceptable two-factor structure.

Based on the factor analysis results, we computed a Stability score and a Plasticity score by averaging items. The descriptive statistics for these scores are shown in Table 4. Paired-sample t-tests were used to compare means on Stability and Plasticity for sibling 1 vs. 2; none was statistically significant. Correlations between the Plasticity and Stability scale scores were computed together for sibling1 and sibling 2, with significance tests adjusted to take into account sibling non-independence (Griffin & Gonzalez, 1995). In all three samples, Stability and Plasticity scores were positively correlated: N2CAP, $r(110) = .27, p < .001$; TRACKS, $r(117) = .21, p < .01$; WRRP, $r(305) = .21, p < .001$.

Score Construct Validity—To examine whether the tester-rated Plasticity and Stability scores corresponded with parents’ reports of child temperament, we computed Pearson correlations together for sibling1 and sibling2. Significance tests were adjusted to take into account sibling non-independence (Griffin & Gonzalez, 1995). Because the same measure of parents’ reports of child temperament was used in the N2CAP and TRACKS samples (i.e., the CCTI) whereas a different measure was used in the WRRP sample (i.e., the CBQ), the correspondence between parent-rated child temperament and tester-rated Stability and Plasticity scores were examined together for the N2CAP and TRACKS samples, and separately for the WRRP sample (Table 5). Because the N2CAP and TRACKS samples were combined, a dummy coded grouping variable was created to control for between-sample sources of variation (Curran & Hussong, 2009). The following pattern of covariation was predicted (and shown as underlined correlations in Table 5): parent-rated surgency/extraversion positively correlated with Plasticity but independent ($r = .0$) of Stability; a composite score of parent-rated effortful control/persistence and negative affectivity/emotionality (reversed) independent of Plasticity but positively correlated with Stability. Of the 4 predicted correlations, 3 were statistically significant and 4 were in the expected direction. Of the 4 correlations for which we did not have predictions, none was significant. Therefore, the data generally matched the predicted pattern.

Sex and Age Differences

Independent sample t-tests were used to investigate potential sex differences in Stability and Plasticity. Of the 12 t-tests conducted (i.e., three studies x two siblings per study x two variables), none was statistically significant after Bonferroni correction adjusting for multiple comparison, suggesting no mean level difference in Stability and Plasticity between boys and girls.

We also examined age differences in the cross-sectional data. We examined linear associations with child age in the N2CAP and WRRP studies by estimating Pearson correlations together for sibling 1 and 2; TRACKS had no variance in child age so those data were not included in this analysis. Significance tests were adjusted to take into account sibling non-independence (Griffin & Gonzalez, 1995). Plasticity was lower ($r = -.10$ to $-.23$, $p < .05$) and Stability was higher ($r = .21$ to $.24$, $p < .01$) among older children.

Sibling Similarity and Behavioral Genetic Analyses

For subsequent analyses, the N2CAP sample was restricted to include only same-sex pairs that also were less than three years apart in age ($n = 36$ pairs). This was done in order to make the adoptive sibling sample as comparable as possible to the same-sex twins in the other two studies, for the purposes of estimating sibling similarity and behavioral genetic parameters. Sibling intra-class correlations are shown in Table 6. Correlations were moderate and significant for the twin samples except for a nonsignificant correlation for Stability among the DZ twins in the TRACKS study. Correlations were modest and nonsignificant for adoptive siblings. Visual inspection of the sibling correlations showed a pattern of $MZ\ r > DZ\ r > Adoptive\ sibling\ r$, with the exception that the correlation in stability between DZ twins in the TRACKS was lower than that between adoptive siblings in the N2CAP. This general pattern suggested the presence of some genetic variance.

We used SEM to examine genetic and environmental variance in Stability and Plasticity. In these models, latent additive genetic (A), shared environmental (C), and nonshared environmental (E) variance together added up to the observed phenotypic variance (Figure 2). The correlation for additive genetic similarity between sibling 1 and sibling 2 was set at 1, .5, and 0 for MZ twin, DZ twin, and adoptive siblings, respectively. The correlations for shared and nonshared environmental variance between sibling 1 and sibling 2 were set as 1 and 0, respectively, for all MZ, DZ, and adoptive siblings. With MX (Neale, 1997), all the models (i.e., ACE model) were fit to the observed sibling variance/covariance matrices using maximum likelihood model-fitting procedure. Fit indices and parameter estimates for the ACE models are presented in Table 7. Model fit was good, with $p > .05$ for χ^2 and $RMSEA = .00$. Individual differences in Stability and Plasticity both included modest additive genetic variance ($a = .10$ to $.36$), modest shared environmental variance ($c = .27$ to $.47$), and modest nonshared environmental variance ($e = .37$ to $.43$), although the modest genetic variance estimate for Stability included 0 within its 95% confidence interval.

Discussion

In recent years, there is increasing interest in examining whether broader and more general personality traits than the Big Five exist, and what their functional roles are (e.g., DeYoung, 2006; Saucier, 2003). Evidence supporting the presence of two broad personality factors (i.e., Stability/Morality & Plasticity/Dynamism) has been obtained using multiple methodologies and informants. In the current study, we followed the reasoning of Caprara et al. (1997) that observers, when facing a large amount of information about an individual's behavior, will use an efficient strategy via more general personality-perception schema to describe another person's behavior. We argue that when direct behavioral observations are used, broad personality factors such as Stability and Plasticity are very likely to emerge. In addition, we extended this literature by testing for the presence of these two broad personality meta-traits in childhood and by examining genetic and nongenetic sources of variance.

What are Stability and Plasticity?

Using EFA in one sample and CFA in another two samples, observer's ratings yielded a clear two factor structure that corresponded well with the results from previous studies (i.e., Stability and Plasticity), suggesting that individual differences in the two broad personality factors are observable by a stranger observer after a few hours of interaction with the target. Rather than method artifacts or merely replication of two of the Big Five factors, we argue that the two-factor structure yielded by stranger observer's ratings reflect meaningful broad personality factors based on the following two reasons. First, although the items were limited in number, they covered a wide range of behaviors that spanned all five personality domains. For example, indicators of the Stability factor included negative affect and frustration (neuroticism), attention and persistence (conscientiousness), and cooperation and orientation to observer (agreeableness); indicators of the Plasticity factor included positive affect and social engagement (extraversion) and exploration (openness to experience). Second, a comparison between observer-rated Stability and Plasticity and standardized parents' ratings of child temperament on closely related dimensions (Table 4) indicated satisfactory correspondence and discrimination of scales across informants. Therefore, the two factors obtained from stranger observer's ratings at least semantically resembled the two general factors obtained in previous studies (Caprara et al., 1997; DeYoung, 2006; Saucier et al., 2005; Slobodskaya, 2011), and captured broader personality characteristics that could not be described by any single one of the Big Five personality traits.

In addition, the current study extended previous literatures by testing the early emergence of these two broad personality factors in children, and results suggested that individual differences in Stability and Plasticity factors were already evident by around 3.5 years of age, consistent with the result from a previous study using parent- and self-report of personality (Slobodskaya, 2011).

Cross-sectional comparisons showed that Plasticity was lower and Stability was higher as a function of child age. This is consistent with a prior study of longitudinal change in the Big Five factors in childhood and adolescence (Lamb, Chuang, Wessels, Broberg, & Hwang, 2002). That study showed that, from 6 to 15 years of age, mean levels of agreeableness and

conscientiousness increased over time, but emotion stability did not show mean level change (these three factors being the components of Stability). In contrast, mean levels of extraversion and openness decreased over time (these two factors being the components of Plasticity). However, results from other studies have been mixed. For example, De Fruyt et al. (2006) found no mean level change over three years in childhood (De Fruyt et al., 2006) and another study found nonlinear changes from late childhood through adolescence that depended on specific facets in question (Soto, John, Gosling, & Potter, 2011). The different results may suggest that mean level shifts that reflect biosocial changes and challenges across childhood and adolescence may be better represented by lower-level facets. In contrast, broader meta-traits may best capture developmental continuities that reflect gradual improvements in psychosocial adjustment.

The existence and early emergence of the two broad personality factors potentially reflects a co-evolution process of biological factors and human society in which complex individual differences are aligned along the two main themes of social living: socialization and personal growth (Digman, 1997). Consistent views can be found in other theories concerning the biological mechanisms underlying personality. These theories have argued for the evolution of specialized neurological systems that contribute to the two broad aspects of personality (DeYoung et al., 2002) and other related broad dimensions of personality (Elliot & Thrash, 2010; Gray, 2000).

Genetic and Environmental Influences

In the current study, we also estimated additive genetic and nongenetic sources of variance in the broad personality factors using a behavioral genetic design. Individual differences in observed Plasticity were explained by nearly equal proportions of additive genetic, shared environmental, and nonshared environmental variance. Stability was mainly accounted for by shared and nonshared environmental variance. The current results were distinct from previous findings in two ways. First, genetic variance in the current study was lower than genetic variance reported in previous studies (.4 – .7 range). Second, one-third of the variance in Plasticity and one half of the variance in Stability was shared environmental variance, in contrast to negligible amounts of this kind of variance in prior studies (Jang et al, 2006). The unique results may reflect the specific methodological features of the current study. Observers' ratings of behavior based on limited periods of observation under constrained circumstances (as was done in the current study) often yield individual difference scores that have less heritability and more shared environmental variance compared to self- or parent-report measures, due to systematic situational influences (Borkenau, Riemann, Angleitner, & Spinath, 2001; Deater-Deckard, 2000; Riese, 1990).

The use of observer's ratings is a strength that enabled us to examine issues that have not been addressed in previous studies of childhood. For example, no previous study has examined the broader personality factors in children using observer's ratings despite the fact that observation is such a widely used method to measure young children's temperament and personality. In addition, new perspectives of stranger observers on child behaviors may bring in additional understanding of the nature and interpretation of these broader personality factors (Caprara et al., 1997). However, there are also limitations to the current

study. Each child was rated only by one observer, which meant we were not able to estimate inter-rater reliability. In addition, because we combined samples for some analyses, it would be ideal to not only estimate inter-rater reliability, but also examine the comparability of the inter-rater reliabilities across the three samples. Another limitation is that, as an initial attempt to explore whether general personality factors could be captured by stranger observers upon observing the targets for a limited period of time, the observational instrument available in our archived datasets contained items that were relatively limited in number and the situation was constrained by the parameters of a “testing” session, and therefore, they may not cover the full range of personality. For example, in such a novel and challenging testing situation, positive affect might be potentially under-represented in children who are fearful and inhibited in particular. The fact that the “fearfulness” item had double loadings on both factors may suggest that such bias is potentially present in the current study. To further this investigation and minimize such a bias, future studies should apply observational instrument that could measure personality characteristics more comprehensively and that is less constrained by situations. Finally, because the situation was potentially constrained by the parameters of a “testing” session, this may have yielded behavioral genetic variance estimates in the broad personality factors that would not generalize to observers’ ratings in other situations or settings, or based on ratings from knowledgeable informants such as parents and peers. In future, it will be essential to purposefully include multi-situation and multi-observer rating schemes into genetically sensitive study designs to address these limitations.

In summary, a broad two-factor (Stability and Plasticity) personality structure was found for children based on observer’s ratings. Most of the sibling similarity in Plasticity was due to additive genetic influences, whereas most sibling similarity in Stability was attributable to shared environmental factors. It is argued that individual differences in the two broad personality factors can be captured by stranger observers after meeting the targets for only a few hours. More importantly, individual differences in the Stability and Plasticity factors were already present by 3.5 years of age. Future studies are needed to explore the biological substrates supporting such an organization of personality (e.g. neurotransmitters, candidate genes). In addition, more longitudinal studies spanning even wider age ranges are needed to obtain a clearer understanding of how these broad personality traits develop throughout the entire life span.

Acknowledgments

We wish to thank the participants, research staff, and funding agencies. N2CAP was supported by a grant from the Society for the Psychological Study of Social Issues, the Kathryn Rudolph Memorial Research Fund of the University of Oregon, and National Science Foundation grants BCS-9907860, BCS-9907811, and BCS-0196511. WRRP was supported by NICHD grant HD38075 and NICHD/OSERS grant HD46167. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institute of Child Health And Human Development or the National Institutes of Health.

References

- Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. 1974; 19:716–723.
- Ashton MC, Lee K, Goldberg LR, De Vries RE. Higher-order factors of personality: Do they exist? *Personality and Social Psychology Review*. 2009; 13(2):79–91. [PubMed: 19458345]

- Barbaranelli C, Caprara GV, Rabasca A, Pastorelli C. A questionnaire for measuring the Big Five in late childhood. *Personality and Individual Differences*. 2003; 34:645–664.
- Bayley, N. Bayley Scales of Infant Development. New York: Psychological Corp; 1969.
- Biesanz JC, West SG. Towards understanding assessments of the Big Five: Multitrait-multimethod analyses of convergent and discriminant validity across measurement occasion and type of observer. *Journal of Personality*. 2004; 72:845–876. [PubMed: 15210019]
- Borkenau P, Riemann R, Angleitner A, Spinath FM. Genetic and environmental influences on observed personality: Evidence from the German observational study of adult twins. *Journal of Personality and Social Psychology*. 2001; 80(4):655–668. [PubMed: 11316228]
- Caprara GV, Barbaranelli C, Zimbardo PG. Politicians' uniquely simple personalities. *Nature*. 1997; 385(6):493.
- Caspi A. The child is father of the man: Personality continuities from childhood to adulthood. *Journal of Personality and Social Psychology*. 2000; 78(1):158–172. [PubMed: 10653512]
- Caspi A, Roberts BW, Shiner RL. Personality development: Stability and change. *Annual Review of Psychology*. 2005; 56:453–484.
- Costa PT Jr, McCrae RR. Reply to Eysenck. *Personality and Individual Differences*. 1992; 13:861–865.
- Curran PJ, Hussong AM. Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods*. 2009; 14(2):81–100. [PubMed: 19485623]
- Deater-Deckard K. Parenting and child behavioral adjustment in early childhood: A quantitative genetic approach to studying family processes. *Child Development*. 2000; 71(2):468–484. [PubMed: 10834478]
- Deater-Deckard K, Petrill SA. Parent-child mutuality and child behavior problems: Gene-environment processes. *Journal of Child Psychology and Psychiatry*. 2004; 45:1171–1179. [PubMed: 15257673]
- De Fruyt F, Mervielde I, Hoekstra HA, Rolland J. Assessing adolescents' personality with the NEO PI-R. *Assessment*. 2000; 7(4):329–345. [PubMed: 11172584]
- De Fruyt F, Bartels M, Van Leeuwen KG, De Clercq B, Decuyper M, Mervielde I. Five types of personality continuity in childhood and adolescence. *Journal of Personality and Social Psychology*. 2006; 91(3):538–552. [PubMed: 16938036]
- DeYoung CG. Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology*. 2006; 91(6):1138–1151. [PubMed: 17144770]
- DeYoung CG, Peterson JB, Higgins DM. Higher-order factors of the Big Five predict conformity: Are these neuroses of health? *Personality and Individual Differences*. 2002; 33:533–552.
- Digman JM. Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*. 1997; 73:1246–1256. [PubMed: 9418278]
- Eisenberg N. Emotion, regulation, and moral development. *Annual Review of Psychology*. 2000; 51:665–697.
- Elliot AJ, Thrash TM. Approach and avoidance temperament as basic dimensions of personality. *Journal of Personality*. 2010; 78(3):865–906. [PubMed: 20573129]
- Freedman DG, Keller B. Inheritance of behavior in infants. *Science*. 1963; 140:196–198. [PubMed: 13959419]
- Friesen CA, Kammrath LK. What it pays to know about a close other: The value of if-then personality knowledge in close relationships. *Psychological Science*. 2011; 22(5):567–571. [PubMed: 21478474]
- Goldberg LR. The structure of phenotypic personality traits. *American Psychologist*. 1993; 48:26–34. [PubMed: 8427480]
- Goldberg LR, Somer O. The hierarchical structure of common Turkish person-descriptive adjectives. *European Journal of Personality*. 2000; 14:497–531.
- Goldsmith HH, Gottesman II. Origins of variation in behavioral style: A longitudinal study of temperament in young twins. *Child Development*. 1981; 52:91–103. [PubMed: 7195330]
- Gray, JA.; McNaughton, N. *The neuropsychology of anxiety*. Oxford, UK: Oxford University Press; 2000.

- Griffin D, Gonzalez R. Correlational analysis of dyad-level data in the exchangeable case. *Psychological Bulletin*. 1995; 118(3):430–439.
- Jang KL, Livesley WJ, Ando J, Yamagata S, Suzuki A, Angleitner A, Spinath F. Behavioral genetics of the higher-order factors of the Big Five. *Personality and Individual Differences*. 2006; 41:261–272.
- Lamb ME, Chuang SS, Wessels H, Broberg AG, Hwang CP. Emergence and construct validation of the Big Five factors in early childhood: A longitudinal analysis of their ontogeny in Sweden. *Child Development*. 2002; 73(5):1517–1524. [PubMed: 12361316]
- Matheny AP, Dolan AM, Wilson RS. Twins: Within-pair similarity on Bayley's Infant Behavior Record. *Journal of Genetic Psychology*. 1976; 128:263–270. [PubMed: 945328]
- McCrae, RR.; Costa, PT, Jr. A Five-Factor theory of personality. In: Pervin, LA.; John, OP., editors. *Handbook of personality: Theory and research*. 2. New York, NY: Guilford Press; 1999. p. 139-153.
- McCrae, RR.; Costa, PT, Jr. Empirical and theoretical status of the five-factor model of personality traits. In: Boyle, GJ.; Matthews, G.; Saklofske, DH., editors. *The SAGE handbook of personality theory and assessment*, vol.1, *Personality theories and models*. Thousand Oaks, CA: Sage Publications; 2008. p. 273-294.
- McCrae RR, Costa PT Jr, Ostendorf F, Angleitner A, Hrebickova M, Avia MD, Smith PB. Nature over nurture: Temperament, personality and lifespan development. *Journal of Personality and Social Psychology*. 2000; 78:173–186. [PubMed: 10653513]
- Musek J. A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality*. 2007; 41:1213–1233.
- Neal, MC. *Mx: Statistical modeling*. Richmond, VA: Department of Psychiatry, Medical College of Virginia; 1997.
- O'Connor BP. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*. 2000; 32(3):396–402.
- Olsen JA, Kenny DA. Structural equation modeling with interchangeable dyads. *Psychological Methods*. 2006; 11(2):127–141. [PubMed: 16784334]
- Paulhus DL, John OP. Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*. 1998; 66(6):1025–1060.
- Petrill SA, Deater-Deckard K, Thompson L, DeThorne L, Schatschneider C. Genetic and environmental effects of serial naming and phonological awareness on early reading outcomes. *Journal of Educational Psychology*. 2006; 98:112–121. [PubMed: 19444324]
- Putnam SP, Rothbart MK. Development of short and very short forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment*. 2006; 87(1):103–113.
- Riese ML. Neonatal temperament in monozygotic and dizygotic twin pairs. *Child Development*. 1990; 61:1230–1237. [PubMed: 2209192]
- Rothbart MK. Temperament, development, and personality. *Current Directions in Psychological Science*. 2007; 16(4):207–212.
- Rothbart MK, Ahadi SA, Hershey L, Fisher P. Investigation of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development*. 2001; 72(5):1394–1408. [PubMed: 11699677]
- Rothbart, MK.; Bates, JE. Temperament. In: Damon, W.; Eisenberg, N., editors. *Handbook of child psychology*, vol. 3, *social, emotional, and personality development*. 5. New York: John Wiley & Sons; 1998. p. 105-177.
- Rowe D, Plomin R. Temperament in early childhood. *Journal of Personality Assessment*. 1977; 41(2):150–156. [PubMed: 856967]
- Rushton JP, Bons TA, Ando J, Hur Y, Irwing P, Vernon PA, Barbaranelli C. A general factor of personality from multitrait-multimethod data and cross-national twins. *Twin Research and Human Genetics*. 2009; 12(4):356–365. [PubMed: 19653836]
- Saucier G. Factor structure of English-language personality type-nouns. *Journal of Personality and Social Psychology*. 2003; 85:695–708. [PubMed: 14561123]

- Saucier G, Georgiades S, Tsaousis I, Goldberg LR. The factor structure of Greek personality adjectives. *Journal of Personality and Social Psychology*. 2005; 88(5):856–875. [PubMed: 15898880]
- Shiner R, Caspi A. Personality differences in childhood and adolescence: Measurement, development, and consequences. *Journal of Child Psychology and Psychiatry*. 2003; 44(1):2–32. [PubMed: 12553411]
- Slobodskaya HR. Two superordinate personality factors in childhood. *European Journal of Personality*. 2011; 25(6):453–464.
- Soto CJ, John OP, Gosling SD, Potter J. Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*. 2011; 100(2):330–348. [PubMed: 21171787]
- Steiger JH. Structure model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*. 1990; 25(2):173–180.
- Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*. 1973; 38:1–10.
- Zwick WR, Velicer WF. Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*. 1986; 99(3):432–442.

Figure 1.
Confirmatory factor analysis of a two-factor structure.

Figure 2.
Univariate behavioral genetic model (i.e., ACE model) of Stability and Plasticity.

Table 1

Distribution of Eigenvalues for Correlations of 17 BBR Variables Based on Real Data from N2CAP and Random Data

Eigenvalue Number	Sibling 1	Sibling 2	Random data
1	7.36	6.77	1.86
2	2.56	2.43	1.52
3	1.33	1.07	1.36
4	.97	.98	1.29

Table 2

Means and Standard Deviations of 13 BBR items

	Mean			Std Dev		
	WRRP	TRACKS	N2CAP	WRRP	TRACKS	N2CAP
a: positive affect	4.29	3.13	3.52	.92	.95	1.33
b: negative affect	4.44	4.52	4.33	.85	.76	1.03
c: soothability	4.50	4.35	4.57	.72	.90	.67
d: hypersensitivity	4.74	4.55	4.51	.57	.86	.87
e: energy	4.45	4.22	4.10	.79	.89	1.15
f: adaptation to change	4.78	4.54	4.55	.56	.77	.75
i : exploration	3.96	3.05	3.56	1.02	.98	1.24
j: attention	4.32	3.83	4.20	.81	.74	.86
k: persistence	4.26	3.82	4.20	.88	.90	.94
n: low frustration	4.58	4.63	4.35	.71	.70	.87
o: oriented to observer	4.75	4.32	4.55	.56	.80	.77
p: social engagement	3.98	3.12	3.70	1.07	1.09	1.37
q: cooperation	4.66	4.28	4.59	.71	.82	.68

Table 3Confirmatory Factor Analyses: Factor Loadings (all significant at $p < .001$) and Model Fit

	Item number	Factor loading
Stability	b: low negative affect	.72
	c: soothability	.58
	d: low hypersensitivity	.58
	f: adaptation to change	.67
	j: attention	.73
	k: persistence	.75
	n: low frustration	.59
	o: oriented to observer	.79
	q: cooperation	.87
Plasticity	a: positive affect	.67
	e: energy	.60
	i: exploration	.65
	p: social engagement	.74
Fit Indices:	RMSEA ^a	.068
	TLI ^b	.927

^aRoot mean square error of approximation;^bTucker-Lewis Index.

Table 4

Means (Standard Deviations)of Plasticity and Stability

	<u>Plasticity</u>	<u>Stability</u>
TRACKS sibling 1 (<i>n</i> = 118)	3.37 (.78)	4.26 (.61)
TRACKS sibling 2 (<i>n</i> = 118)	3.40 (.74)	4.37 (.57)
WRRP sibling 1 (<i>n</i> = 309)	4.14 (.73)	4.57 (.52)
WRRP sibling 2 (<i>n</i> = 309)	4.20 (.79)	4.55 (.57)
N2CAP sibling 1 (<i>n</i> = 114)	3.60 (1.06)	4.51 (.57)
N2CAP sibling 2 (<i>n</i> = 111)	3.83 (.95)	4.35 (.64)

Table 5

Pearson Correlations between Tester-Rated Plasticity and Stability and Parent-Rated Temperament

	Plasticity	Stability
WRRP:		
Surgey	<u>.19</u> ***	-.03
Effortful Control & Negative Affectivity (reversed) composite	-.10	<u>.26</u> ***
N2CAP & TRACKS:		
Extraversion composite	<u>.09</u>	-.01
Persistence & Emotionality (reversed) composite	-.02	<u>.12</u> *

Note: Predicted correlations are underlined. Sample sizes for sibling pairs: WRRP (n = 208), N2CAP & TRACKS (n = 182).

**
p < .01,

p < .001.

Table 6

Sibling Intra-Class Correlations for WRRP, TRACKS, and N2CAP Studies

	MZ twins		DZ twins		Adoptive sibs	
	<u>WRRP</u>	<u>TRACKS</u>	<u>WRRP</u>	<u>TRACKS</u>	<u>TRACKS</u>	<u>N2CAP</u>
Plasticity	.63***	.64***	.26**	.41**		.10
Stability	.58***	.55***	.60***	.19		.28

**
p < .01,

p < .001

Table 7

Genetic, Shared Environmental, and Nonshared Environmental Variance (95% confidence intervals) and Model Fit Indices for the ACE Models

	a^2	c^2	e^2	χ^2 (df, p)	RMSEA	AIC
Plasticity	.36 (.23 – .50)	.27 (.15 – .38)	.37 (.30 – .46)	2.67 (9, .98)	.00	–5.20
Stability	.10 (.00 – .22)	.47 (.37 – .56)	.43 (.35 – .53)	16.12 (9, .06)	.00	–1.88

Note: a^2 : additive genetic variance; c^2 : shared environmental variance; e^2 : nonshared environmental variance.