

# Genetic and Environmental Continuity in Personality Development: A Meta-Analysis

Daniel A. Briley and Elliot M. Tucker-Drob  
University of Texas at Austin

The longitudinal stability of personality is low in childhood but increases substantially into adulthood. Theoretical explanations for this trend differ in the emphasis placed on intrinsic maturation and socializing influences. To what extent does the increasing stability of personality result from the continuity and crystallization of genetically influenced individual differences, and to what extent does the increasing stability of life experiences explain increases in personality trait stability? Behavioral genetic studies, which decompose longitudinal stability into sources associated with genetic and environmental variation, can help to address this question. We aggregated effect sizes from 24 longitudinal behavioral genetic studies containing information on a total of 21,057 sibling pairs from 6 types that varied in terms of genetic relatedness and ranged in age from infancy to old age. A combination of linear and nonlinear meta-analytic regression models were used to evaluate age trends in levels of heritability and environmentality, stabilities of genetic and environmental effects, and the contributions of genetic and environmental effects to overall phenotypic stability. Both the genetic and environmental influences on personality increase in stability with age. The contribution of genetic effects to phenotypic stability is moderate in magnitude and relatively constant with age, in part because of small-to-moderate decreases in the heritability of personality over child development that offset increases in genetic stability. In contrast, the contribution of environmental effects to phenotypic stability increases from near zero in early childhood to moderate in adulthood. The life-span trend of increasing phenotypic stability, therefore, predominantly results from environmental mechanisms.

**Keywords:** behavioral genetics, personality, rank-order stability, cumulative continuity, meta-analysis

**Supplemental materials:** <http://dx.doi.org/10.1037/a0037091.supp>

A defining feature of psychological traits is that they are relatively stable across time (Allport, 1937). *Differential stability* refers to the degree to which the relative differences between individuals are preserved over time. This aspect of stability is typically assessed with a test–retest correlation. Personality traits—individual differences in general patterns of thoughts, feelings, and behavior—display some of the highest differential stabilities of all psychological traits (Conley, 1984). Nevertheless, personality is not uniformly stable over development. For instance, an influential meta-analysis of test–retest data from 152 longitu-

dinal studies of personality found increases in 7-year stability coefficients from .3 in early childhood to .6 by early adulthood and to .7 by later adulthood (Roberts & DelVecchio, 2000).

Both genetic and environmental mechanisms may contribute to patterns of increasing personality stability with age. For example, developmental increases in differential stability could result from the cumulative effects of living in a stable environment and the decreasing occurrence of experiencing novel environments with age and/or from the continuous action of the same genes over long periods of time. Although these hypotheses have played prominent roles in theoretical accounts of personality development (Caspi & Roberts, 2001; Fraley & Roberts, 2005; Loehlin, 1992; McCrae et al., 2000), there has, until recently, been surprisingly little work in this area using behavioral genetic methodologies capable of testing for genetic and environmental mediation of differential stability. For instance, Roberts and DelVecchio (2000) cited a single behavioral genetic study in their meta-analysis. This study, conducted by McGue, Bacon, and Lykken (1993), used a sample of twins to estimate that approximately 80% of 10-year personality test–retest correlations were mediated by genetic factors. Roberts and DelVecchio commented at the time that

unfortunately, longitudinal twin studies of personality development are relatively rare, and no other research has replicated McGue et al.'s findings across the life course. Therefore, it is not known whether genetic influence on consistency increases or decreases across the life course. (Roberts & DelVecchio, 2000, p. 4)

---

This article was published Online First June 23, 2014.

Daniel A. Briley and Elliot M. Tucker-Drob, Department of Psychology and Population Research Center, University of Texas at Austin.

Elliot M. Tucker-Drob was supported by National Institutes of Health (NIH) Research Grant R21HD069772. Daniel A. Briley was supported by NIH Training Grant T32HD007081. The Population Research Center at the University of Texas at Austin is supported by NIH Center Grant R24HD042849. We appreciate very helpful feedback and suggestions from Paige Harden and John Loehlin. Daniel A. Briley presented initial results of this meta-analysis at the January 2013 meeting of the Society for Personality and Social Psychology in New Orleans, Louisiana.

Correspondence concerning this article should be addressed to Daniel A. Briley, Department of Psychology, University of Texas at Austin, 108 East Dean Keeton Stop A8000, Austin, TX 78712-0187. E-mail: [daniel.briley@utexas.edu](mailto:daniel.briley@utexas.edu)

Longitudinal behavioral genetic studies of personality provide two complementary sorts of information that are essential for understanding personality development. First, they provide estimates of the magnitudes of genetic and environmental influences on personality (i.e., the proportions of variance in personality attributable to genetic and environmental factors) at each point in time. Second, they provide estimates of the stabilities of genetic and environmental influences across time (i.e., the correlation between genetic or environmental factors at two time points). Importantly, these two sorts of information are statistically and logically independent of one another, and each component can change over the life span. By combining these two sorts of information, one can calculate the extents to which genetic and environmental factors contribute to overall stability. The current project meta-analyzes longitudinal behavioral genetic studies to address three core questions: (a) What are the relative contributions of genes and the environment to personality across the life span? (b) How stable are the genetic and environmental influences on personality across the life span? (c) To what extent do genetic and environmental factors contribute to the increasing stability of personality across the life span? We begin by reviewing influential theoretical models of personality development and discussing how behavioral genetic methodology can discriminate among competing models.

### Theories of Personality Development

Historically, researchers attempting to explain the increasing stability of personality with age have differentially emphasized intrinsic maturational processes versus exogenous/social processes. That endogenous processes underlie personality development has a long tradition within personality psychology (Barenbaum & Winter, 2008). For example, the early theories of psychosexual development (Freud, 1908/1959) focused on internal processes, and Allport (1937, p. 48) identified personality as relying on internal “psychophysical systems.” Some early proposed mechanisms implicated differences in body chemistry (Murray, 1938) and structure (Sheldon, Stevens, & Tucker, 1940). This perhaps intuitive belief has carried on, with modern personality theories based on the relative influence or presence of neurotransmitters (Cloninger, 1998) and the structure of the reticular activation system within the brain (Eysenck & Eysenck, 1985). These works assume that intrinsic, specifically genetic, maturational processes are the leading causes of personality development. However, there are many personality theories that implicate exogenous influences as shaping personality development (D. P. McAdams & Olson, 2010). Erikson (1950/1963) saw personality as developing in relation to success or failures with certain social challenges. Similarly, attachment theories posit that early caregiving experiences have a lasting influence on a multitude of aspects of an individual’s life (Bowlby, 1973).

How have these broad, historical theories of personality development been applied to the empirical trend of increasing trait stability? Although there are a great number of modern theories of personality development (see Mroczek & Little, 2006), the two frameworks that have been most interested in explaining increasing stability are the five-factor theory (FFT; McCrae & Costa, 2008) and social personality models (SPMs). These SPMs have been alternatively called the neosocioanalytic (Roberts & Wood,

2006) or the sociogenomic (Roberts & Jackson, 2008) model of personality, and they are highly connected to social investment theory (Roberts, Wood, & Smith, 2005). Therefore, we refer to this general class of theories that emphasize social inputs into personality development as SPMs. The FFT takes the strong position that personality development is guided by intrinsic processes. SPMs emphasize that genes or intrinsic processes set the stage for personality development. They hold that life-span trends, however, are primarily a function of exogenous forces, most notably the social environment. Importantly, even in the primary proponents’ most polemic work, both sides have acknowledged that genes are a substantial influence on stability (e.g., Roberts, Wood, & Caspi, 2008, p. 384) and that the environment is a likely influence on at least the expression of personality (McCrae et al., 2000, p. 175). Differences emerge in the interpretation of results as being primarily explained by intrinsic maturation processes or by social mechanisms (see, e.g., Costa & McCrae, 2006; Roberts, Walton, & Viechtbauer, 2006). Below, we provide a brief overview of these theories and how they account for the empirical trend of increasing trait stability.

### Intrinsic Maturation

The intrinsic maturation perspective emphasizes the role that genetic effects play in increasing the stability of personality. In a landmark exposition of this perspective, McCrae et al. (2000) concluded that the empirical evidence of “heritability, limited parental influence, structural invariance across cultures and species, and temporal stability all point to the notion that personality traits are more expressions of human biology than products of life experiences” (p. 177). Under this model, personality traits represent basic tendencies of behavior that are solely influenced by biological (i.e., genetic) mechanisms and are “insulated from the direct effects of the environment” (McCrae & Costa, 2008, p. 164). The environment may alter the characteristic expressions of the basic tendencies (e.g., changing cultural norms for what constitutes politeness) or the biological bases of personality traits (e.g., brain damage), but environmental events such as occupational or romantic roles cannot have a direct impact on personality. Additional empirical support for this biologically oriented perspective includes molecular genetic associations with personality (de Moor et al., 2012; Terracciano, Sanna, et al., 2010; Vinkhuyzen et al., 2012; Webb et al., 2012), parallel cross-cultural age trends in population-mean trait levels (McCrae et al., 1999; cf. Bleidorn et al., 2013), and findings of personality change as a result of alterations to the biological bases of personality in terms of extreme trauma (Damasio, Grabowski, Frank, Galaburda, & Damasio, 1994), parasitic infection (Lindová, Příplatová, & Flegr, 2012), surgery (Zhao et al., 2012), and pharmaceutical effects (Tang et al., 2009), to name only a few examples.

On the basis of early work examining the stability of personality across the life span, McCrae and Costa (1994) provocatively claimed that personality traits were relatively fixed after age 30, as the best available evidence at the time indicated that test–retest stability peaked at this point. Of course, personality traits do exhibit change (test–retest stability is never perfect), but in the context of the larger life-span trend, age 30 appeared to be a turning point from rather large gains in stability during adolescence into firm adult personality constructs. Terracciano, Costa,

and McCrae (2006) found evidence supportive of this claim, in that personality stability was uncorrelated with age after age 30 for each Big Five factor and facet. One explanation for the age 30 plateau (described more fully in the next section) is that mature personality traits (i.e., increased dominance, agreeableness, conscientiousness, and emotional stability) are required for adult occupational roles. Therefore, individuals who already possess this constellation of traits might display greater stability at a younger age. In contrast, Terracciano, McCrae, and Costa (2010) found evidence that the plateau of stability at age 30 was unrelated to earlier personality maturity. The authors argued that such insensitivity to normative maturity is strong evidence that intrinsic mechanisms are guiding personality development rather than social pushes.

To explain the role of intrinsic maturation, proponents of the FFT often make an analogy between personality and height, a phenotype that obviously depends on environmental inputs (e.g., nutrition) but follows a developmental course largely determined by intrinsic mechanisms when sufficient environmental resources are available (e.g., genetic influences; McCrae & Costa, 2006). During childhood and adolescence, there are individual differences in terms of growth spurts that reorder the relative ranking of individuals. Therefore, in adolescence, one would expect relatively low differential stability of height. In early adulthood, it is much less common for individuals to be reordered in terms of their height. McCrae and Costa (2006) argued that personality traits can be understood in a similar way. Ultimately, McCrae and Costa (2008) claimed that “the course of personality development is determined by biological maturation, not by life experiences” (p. 167). Thus, the FFT would predict high stability of genetic effects that would most likely peak near age 30. The phenotypic stability of traits would be predicted to be largely mediated by genetic factors, and the increase in phenotypic stability should be strongly tied to the increasing stability of genetic effects. Finally, the FFT would predict that environmental influences are largely unstable and unimportant for increasing phenotypic stability.<sup>1</sup>

## Social Maturation

SPMs affirm the importance of genetic influences for facilitating trait stability but posit environmental forces as having true, causal influences on personality development. For example, the sociogenomic model of personality specifies that the environment has a causal effect on the function of genes. Instead of the biological bases of personality being completely shielded from the environment, the genome “is intrinsically dependent on the environment for activation and maintenance” (Roberts & Jackson, 2008, p. 1528). Beginning in the early 2000s, Roberts and colleagues (Caspi & Roberts, 2001; Caspi, Roberts, & Shiner, 2005; Roberts & Wood, 2006; Roberts et al., 2008) have laid out a growing number of postulated principles and mechanisms for social maturation (there are currently 26<sup>2</sup>) to explain the complex relations between genes, the environment, and personality development. Although other versions of SPMs of this sort do not directly deal with genetic effects to the same extent as the sociogenomic model, each perspective emphasizes the causal nature of socialization.

Most relevant to the current discussion are Roberts and colleagues’ *cumulative continuity principle* and *plasticity principle*. These principles hold that personality displays increasing stability

throughout the life span (past age 30) and that personality is always amenable to change. According to this perspective, personality stability continues to increase throughout the life span because individuals continue to develop their identity and select into environmental niches. Plasticity of personality follows directly from this logic, in that trait development is thought to never be complete and is thus always open to environmental manipulation. Several important life transitions, such as the entry into the work environment or into a romantic relationship, may thus explain the dramatic change in increasing stability around age 30. Stability is thought to increase following adolescence, as the experiences that shape children in high school and college become more consistent. After major life events, such as obtaining mature romantic or occupational roles, the environment has reached a relatively stable point.

Proponents of SPMs have argued that mature personality profiles are needed for many adult social roles and have presented evidence of increased personality stability among individuals with a personality profile marked with agreeableness, conscientiousness, and emotional stability (Donnellan, Conger, & Burzette, 2007; Roberts, Caspi, & Moffitt, 2001; cf. Terracciano, McCrae, & Costa, 2010). Compared to the empirical evidence for intrinsic maturation, findings of social influences on personality are rarer, but supportive results are beginning to accumulate rapidly. For example, personality change has been associated with military service (Jackson, Thoemmes, Jonkmann, Lüdtkke, & Trautwein, 2012), negative life events (Kandler, Bleidorn, Riemann, Angleitner, & Spinath, 2012), macroenvironmental differences (Bleidorn et al., 2013), the timing of important life events (Bleidorn, 2012), the adoption of social roles (Bogg, Finn, & Monsey, 2012), cognitive training (Jackson, Hill, Payne, Roberts, & Stine-Morrow, 2012), mindfulness training (Krasner et al., 2009), and international travel (Zimmermann & Neyer, 2013). Importantly, each of these studies took steps to rule out alternative selection explanations for personality change.

Roberts (2009) argued that a primary goal of parenting is the socialization of personality, writing, “the goal of parenting and schooling is to arm children with skills, abilities, and character structures . . . we attempt to provide children with personality traits” (p. 138). In explicating how this process occurs, Roberts pointed to consistent, bottom-up effects of the environment on temporary states that “take on a significant causal and mediational role as [states] account for the path through which prolonged environmental effects will change neuroanatomical structures or gene expression, and thus change personality traits” (Roberts, 2009, p. 141). The example given in the current context is of a professor who increases course organization, and by extension generalized behavioral organization across situations, because of

<sup>1</sup> In the most recent, complete exposition of the FFT, McCrae and Costa (2008, p. 163) included a new pathway of influence from external influences to the biological bases of personality (which can then influence personality traits). Through this pathway, it is possible that there could be stable, environmental influences on personality mediated through neurological change, but McCrae and Costa argued that this effect is “outside the confines of personality proper” (McCrae & Costa, 2008, p. 168). Therefore, a charitable interpretation of the FFT would be that some environmental effects may exist but that major life-span trends should still be predominantly driven by genetic factors.

<sup>2</sup> This number is based on a count taken from Roberts et al. (2008).



the continued approval of students for this type of behavior. The student approval generates satisfaction in a social role that acts to reinforce organizational skills. Because these types of environmental experiences are able to get under the skin of individuals, “DNA sequences are not the simple, unchanging causal mechanisms depicted in typical biological personality models,” and “environments can and do affect physiological systems, even one as basic as DNA” (Roberts & Jackson, 2008, p. 1532). Therefore, based on this perspective, one would expect significant, direct environmental influences on personality development. Environmental stability would be predicted to be low in early ages but to increase substantially throughout the entire life span, with some decline in old age when social environments change considerably. Genetic effects might be predicted to be fairly unstable, as the sociogenomic model emphasizes that environments are capable of altering gene expression across the life span. Increases in phenotypic stability would largely be explained by increasingly stable contributions from the environment, with genetic contributions remaining largely constant across the life span.

### Considering Gene–Environment Interaction and Correlation

Although behavioral genetic perspectives have historically emphasized what might be described as strong genetic determinism, current behavioral genetic thinking emphasizes dynamic and interactive processes by which genetic and environmental influences combine to influence behavioral development (Johnson, Penke, & Spinath, 2011; Krueger & Johnson, 2008). *Gene × Environment interaction* occurs when heritable variation in a phenotype is differentially expressed in different environmental contexts (Johnson, 2007; Krueger, Johnson, & Kling, 2006). *Gene × Environment interaction* can lead to age-related changes in heritability as a result of changing environmental contexts with age, such as entry into school, the workforce, or a long-term pair bond. Age-related changes in heritability can also result from biological changes associated with development, such as puberty. *Gene–environment correlation* refers to the nonrandom exposure of individuals to environmental experiences as functions of their genotypes (Plomin, DeFries, & Loehlin, 1977; Scarr & McCartney, 1983). In the case of *passive* gene–environment correlation, parental genotypes both influence the rearing experiences that parents provide to their offspring and are inherited by those offspring. In *active* and *evocative* gene–environment correlation, individuals select and evoke different environments on the basis of their genetically influenced preferences, motivations, and traits. These experiences may in turn affect psychological development.

Empirical evidence for the relevance of *Gene × Environment interaction* and *gene–environment correlation* for personality development comes from both molecular and quantitative genetic studies. For instance, the link between received parenting and personality development differs as a function of child candidate genes implicated in neuromodulation (Bakermans-Kranenburg & van IJzendoorn, 2006; Belsky & Beaver, 2011; Kaufman et al., 2006; Sheese, Voelker, Rothbart, & Posner, 2007; Smith et al., 2012; but see also Luijk et al., 2011). Similarly, research using twin and family methods has found that the magnitudes of genetic influences on personality differ across environmental contexts (Ganiban, Ulbricht, Saudino, Reiss, & Neiderhiser, 2011; Krueger,

South, Johnson, & Iacono, 2008). These *Gene × Environment interaction* results indicate that genetic influences on personality are differentially expressed in certain environmental contexts.

With respect to *gene–environment correlation*, candidate genes have been linked to a number of seemingly environmental outcomes relevant for personality development, including parent, romantic, and peer relationships (Burt, 2008; Dick et al., 2006; Lucht et al., 2006). Similarly, twin and family studies have consistently found that environments are heritable (Avinun & Knafo, 2014; Kendler & Baker, 2007; Klahr & Burt, 2014; Riemann, Kandler, & Bleidorn, 2012; Saudino & Plomin, 1997; Sturaro, Denissen, van Aken, & Asendorpf, 2008). This result means that family members who are more genetically similar are more likely to experience similar environments. As Plomin, Reiss, Hetherington, and Howe (1994, p. 32) commented, such results seem counterintuitive because “environments have no DNA and thus cannot show genetic effects.” However, a viable explanation for this seemingly counterintuitive result is that genes influence individual dispositions toward engaging with different sorts of environments. For example, T. A. McAdams, Gregory, and Eley (2013) found that the genetic influences on exposure to negative life events and parenting could be accounted for by the genetic influences on personality characteristics.

The effects of *Gene × Environment interaction* may have implications for personality development and provide insights into the types of life-span trends to expect for behavioral genetic estimates. Early in the life span, children may differentially respond to the common environment provided by their caregivers on the basis of genotypes. This type of *Gene × Shared Environment interaction* results in environmental effects becoming coupled with genetic variation. Genetically similar individuals respond to the environment similarly and become *more* psychologically similar than individuals who share fewer genes. In other words, the effects of *Gene × Shared Environment interaction* contribute to the estimate of genetic influence when not explicitly modeled (Purcell, 2002). Under such circumstances, genetic stability will be high when the same sorts of environments either recur over development or have a lasting impact. Moreover, if such effects compound over time, heritability would be expected to increase with age. However, it may also be the case that early *Gene × Shared Environment interactions* for personality development are fleeting and give way to less genetically dependent environmental influence (e.g., Conley, 1984; Wolf & Weissing, 2012). If the effects of *Gene × Shared Environment interaction* fade over time, then it would be expected that initially high levels of heritability would decline with age. Genetic stability would be low in early life while the effects of environmentally dependent gene expression weaken, but high later in the development.

As individuals leave the home environment in adolescence and adulthood, it is more likely that unique environments, as opposed to those shared with their siblings, will have an impact on personality (Bouchard & Loehlin, 2001). *Gene × Nonshared Environment interaction* has the effect of making genetically similar individuals *less* similar, and therefore genetic effects become tied to unique environmental variation. In other words, the effects of *Gene × Nonshared Environment interaction* contribute to the estimate of nonshared environmental influences when not explicitly modeled (Purcell, 2002). In this case, nonshared environmental stability would be expected to rise as individuals mature and leave the

shared home environment. Again, such interactions may be recurring, lasting, or fleeting. If they are recurring or lasting, this would result in stable nonshared environmental influences. If they are fleeting, this would result in unstable nonshared environmental influences.

Gene–environment correlation may also have implications for personality development. Passive gene–environment correlation leads to increased sibling similarity, regardless of genetic similarity between siblings. In other words, the effects of passive gene–environment correlation contribute to the estimate of shared environmental influence when not explicitly modeled (Purcell, 2002). Active and evocative gene–environment correlations result in non-random environmental experiences, such that environmental effects become tied to genotypic differences. In other words, the effects of active and evocative gene–environment correlation contribute to the estimate of genetic influence when not explicitly modeled (Purcell, 2002).

The effects of gene–environment correlation may also differ across development. Passive gene–environment correlation would be expected to play the largest role early in development when children are still located in the home environment. As children mature, they have increasing autonomy to select experiences based on their genetically influenced dispositions and a greater variety of possible environments from which to select experiences (Scarr & McCartney, 1983). Therefore, patterns of active and evocative gene–environment correlation would be expected to manifest in increasing heritability and, if they are recurrent or lasting, increasing stability of genetic effects with age. Proponents of developmental models based on gene–environment correlation (e.g., Dickens & Flynn, 2001) argue that environments encountered by happenstance are far less likely to recur with sufficient enough frequency to have prominent effects on psychological development. Alternatively, environments systematically selected and evoked on the basis of gene–environment correlation are likely to recur routinely, such that their effects on psychological development are pronounced. Thus, the stability of nonshared environmental effects on personality, which only includes effects of environmental experiences that are not systematically related to genotype, might be expected to be relatively lower than that of genetic effects on personality, which additionally includes effects of environmental experiences that are systematically related to genotype. However, as discussed earlier, transitions to increasingly stable life roles that occur with age would be expected to result in increasing nonshared environmental stability with age.

### Behavioral Genetic Models of Personality Stability

Simple behavioral genetic findings are typically referenced by personality researchers to highlight that personality traits are partially heritable. However, as discussed above, behavioral genetic research also provides insight into the stability of genetic and environmental influences in addition to the relative contributions of genetic and environmental mechanisms to trait stability. In this section, we provide an overview of behavioral genetic models for both single-occasion and longitudinal data and formally define what we mean by stability of genetic and environmental effects and genetic and environmental contributions to stability.

Classical behavioral genetic models use data collected from similarly aged siblings with varying degrees of genetic relatedness

to decompose variation in a trait into that associated with genetic and environmental factors (Neale & Cardon, 1992). The typical behavioral genetic approach, for instance, uses data from monozygotic and dizygotic twins reared together and leverages the knowledge that monozygotic twins share nearly identical genetic material and dizygotic twins share approximately half of segregating genetic material on average. Genetic influences are then inferred to operate on a trait if monozygotic twins resemble one another more on that trait than dizygotic twins. Variance in the trait is decomposed into an additive genetic component (A), a shared environmental component (C), and a nonshared environmental component (E). The E component represents environmental influences that cause siblings to be less similar to one another and includes measurement error (which is by definition uncorrelated across siblings). The C component represents environmental experiences that cause siblings raised within the same home to be more similar to each other. However, a large body of literature indicates that personality traits are only rarely influenced by this class of effects (Bouchard & Loehlin, 2001; but see also Krueger et al., 2008). As such, we focus our attention on AE models.

A path diagram for a univariate AE model for single-occasion data is presented in Figure 1. Variance in this trait is decomposed into latent A and E factors (represented as circles). The correlation between A factors is fixed to specific values depending on genetic relatedness of the siblings. For monozygotic twins (who have

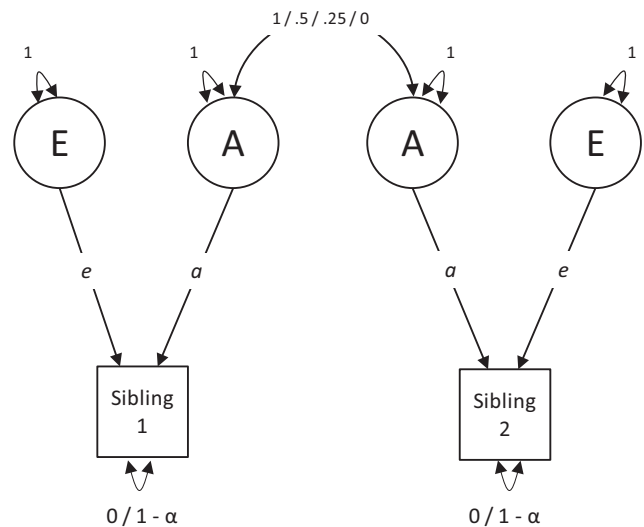


Figure 1. Univariate behavioral genetic model for single-occasion data that decomposes variance in a trait (indicated as a square) into that which is due to latent genetic (A) and environmental (E) components (indicated as circles). The correlation between genetic factors is specified for each group depending on the known genetic association between siblings. The label placed on this parameter in the figure is for each sibling type found in the current study, namely, monozygotic twins (1), dizygotic twins (0.5), half-siblings (0.25), and unrelated siblings (0). When the environmental component is corrected for measurement error, the residual variance of the trait is set to equal  $1 - \text{reliability}$ . Parameters that share the same label are constrained to be equal. When the outcome is standardized before analysis, as is the case in the current analysis, the squared  $a$  and  $e$  parameters represent the proportion of variance in the trait attributable to A and E, respectively.

nearly identical genotypes), the correlation between the A factors of the siblings is fixed to 1. For dizygotic twins (who share approximately half of segregating genetic material on average), this correlation is fixed to .5. For half-siblings (who share approximately one quarter of segregating genetic material on average), the correlation between A factors is fixed to .25. The E factors are unique for each sibling and are always uncorrelated no matter what sibling type is represented. If the outcome is standardized before analysis, squaring the paths labeled  $a$  and  $e$  yields the proportion of variance in the phenotype attributable to the latent effects of A and E.

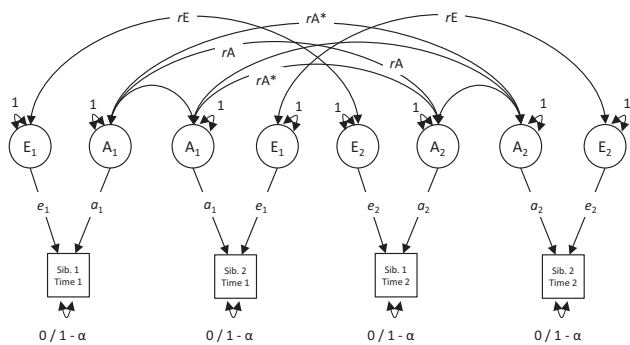
Longitudinal extensions of this methodology decompose both time-point-specific variation and variation that is stable across time points into genetic and environmental components. Stable genetic or environmental influences include effects that are present at two time points and account for variance in personality. The longitudinal correlated factors model presented in Figure 2 is an example of this type of analysis. In addition to the contribution of A and E at an initial time point ( $a_1$  and  $e_1$ ), a second time point is included, and  $a_2$  and  $e_2$  represent the genetic and environmental influences at the second time point. The latent A and E factors at each time point are correlated ( $rA$  and  $rE$ ), indicating whether the same or different genetic and environmental effects are present at the two time points. This analysis derives information from cross-sibling cross-time correlations. If one sibling's initial personality predicts the other sibling's later personality to a greater degree for pairs that are more genetically related, then this result would be indicative of stable genetic influence across time points. Note that, by definition, the environmental influences are uncorrelated across

siblings. They are, however, allowed to correlate across time points within individuals (e.g., E at Time 1 for Sibling 1 is only correlated with E at Time 2 for Sibling 1). Genetic effects, however, are correlated across siblings and time points to the extent that the siblings share genetic material. This specification is denoted in the model as  $rA^*$ . For example, the cross-twin cross-time genetic correlation for monozygotic twins would simply be the genetic correlation ( $rA^* = 1 \times rA$ ), but for dizygotic twins, this pathway would be constrained to be equal to half of the genetic correlation ( $rA^* = 0.5 \times rA$ ). Because the same individual retains the same genotype across time points, the within-sibling cross-time pathway is equal to  $rA$ .

Importantly, genetic stability ( $rA$ ) is not directly tied to the magnitude of genetic influence ( $a$ ) on the trait. A highly heritable phenotype may be influenced by the same genetic factors at two time points, and similarly, a largely environmentally influenced phenotype may still be influenced (to a minor degree) by the same genetic factors at two time points. Despite differences in heritability across these two scenarios, genetic stability may be the same. Likewise, a phenotype may reflect large or minimal environmental effects, but only if these influences persist through time will environmental stability be high. The magnitude and stability of genetic and environmental effects provide distinct pieces of information.

Understanding the genetic and environmental contributions to overall phenotypic stability requires combining information concerning time-specific magnitude and cross-time stability of genetic and environmental factors. For example, a phenotype may be highly heritable at two time points (e.g., childhood and late adolescence), but if genetic stability is low across this period (e.g., as a result of gene activation and deactivation related to puberty), then genetic factors would not contribute to the stability of the phenotype. Similarly, a phenotype may be highly environmental at two time points, but the environmental influences (e.g., peer groups) may change during transitions to new environments (e.g., going from high school to college). The contributions of genetic and environmental effects to stability are calculated by multiplying the time-specific factor loadings by the associated latent variable correlation. For example,  $a_1 \times rA \times a_2$  yields the genetic contribution to stability. Summing the contribution from A and E recaptures the standard test-retest correlation (i.e., phenotypic stability =  $[a_1 \times rA \times a_2] + [e_1 \times rE \times e_2]$ ). Phenotypic stability results from genetic and environmental influences that persist through time.

Two clarifying points are important to mention. First, this analysis uses the same information but is slightly different from another commonly estimated longitudinal association, bivariate heritability. Rather than being reported in terms of raw units, bivariate heritability represents the proportion of a phenotypic correlation that is due to genetic effects. For example, if we observed that a trait displayed a phenotypic stability of .6 across time, a heritability of .5 at each time point, and a genetic correlation of .8, the contribution of genetic influences to phenotypic stability would be .40 (i.e.,  $\sqrt{.5} \times .8 \times \sqrt{.5}$ ), and the bivariate heritability would be .67 (i.e.,  $[\sqrt{.5} \times .8 \times \sqrt{.5}]/.6$ ). The contribution to phenotypic stability has the useful property of being placed on the meaningful metric of phenotypic stability (which changes in magnitude across development) rather than being a proportion. Second, this type of analysis refers to the *effective*



**Figure 2.** Longitudinal correlated factors model that decomposes variance in repeated assessments of personality into that due to time-specific genetic (A) and environmental (E) components, as well as the temporal stability of the genetic ( $rA$ ) and environmental ( $rE$ ) components. Interpretation of the parameters and modifications across sibling groups is the same as in Figure 1 with two notable exceptions. First, that the within-time correlation between genetic factors differs by sibling type has been removed from this figure but is still essential for the model. Second and relatedly, the cross-time cross-sibling genetic correlation (i.e., the correlation between Sibling 1's A factor at Time 1 and Sibling 2's A factor at Time 2) is specified to differ by sibling type such that the expected correlation is scaled relative to the amount of shared genetic material between the siblings (i.e., multiplied by 1 for monozygotic twins and .5 for dizygotic twins, etc.). These genetic correlations have been marked with an asterisk. The within-sibling cross-time genetic correlation is equal to  $rA$ . Cross-sibling cross-time environmental influences are constrained to zero by definition. Sib. = sibling.

rather than *objective* stability of genetic and environmental effects (Turkheimer & Waldron, 2000). Objectively, one's genotype (e.g., presence or absence of a specific polymorphism) or an environmental input (e.g., presence or absence of a spouse) may be the same at two time points, but the effective influence may be different. Effectively stable genes or environments are those that account for variance in the trait at both time points.

## Methodological Considerations and Hypothesized Moderators

### Error Correction

Ferguson (2010) conducted an updated meta-analysis largely confirming the findings of Roberts and DelVecchio (2000). His analysis emphasized the importance of correcting stability coefficients for measurement error of the indicators. Measurement error is by definition unsystematic and therefore unable to be correlated across time points. Measurement error attenuates the true amount of stable variation in personality. Ferguson's estimated stability coefficients peaked and plateaued in early adulthood, but at a much higher level that approximated a "set in plaster" limitation of personality change (e.g.,  $r_{12} = .94$  for ages 28–34). Measurement error may pose special problems for research on the stability of personality if assessment inventories differ in reliability across different age groups.

Behavioral genetic methods conceptualize measurement error in a slightly different manner. Because measurement error is unsystematic, it serves to make siblings different from one another and is therefore a component of the nonshared environment. Typically, the variance of a measured outcome in a twin model is assumed to be 0 as the A and E latent factors represent the total decomposition of this variance. To correct the E component for a known amount of measurement error, the phenotype can be specified to have a set amount of residual (error) variance not explained by A or E. Thus, the remaining E variance represents true environmental effects on the outcome. Figures 1 and 2 explicitly depict this specification by showing that the phenotype's residual variance is 0 in the standard model (this specification is often left out of depictions of behavioral genetic models) and as  $1 - \text{reliability}$  for a model in which measurement error is corrected.

The correction can be interpreted in standardized or unstandardized terms. In standardized terms, the proportion of variance attributable to A and E is rescaled relative to the corrected phenotypic variance; this correction increases the apparent influence of A compared to a model that does not correct for measurement error. The variance attributable to A and E sums to the total true variance in the outcome. Alternatively, one can interpret the unstandardized parameters; the estimate of E is lowered relative to a model that does not correct for measurement error, but the estimate of A is unaltered. We chose to interpret unstandardized parameter estimates. In the context of the correlated factors model, correcting measurement error has the effect of reducing  $e_1$  and  $e_2$  by an amount equal to the measurement error and increasing  $rE$  to indicate the amount of true environmental stability removed of idiosyncratic measurement error. However, the contribution of the environment to phenotypic stability (i.e.,  $e_1 \times rE \times e_2$ ) would remain unchanged, and therefore, the estimated phenotypic stabil-

ity would also remain unchanged. Because increases in  $rE$  are counterweighted by decreases in  $e_1$  and  $e_2$ , correcting for measurement error does not affect aggregates of these parameters. With this nuance in mind, we corrected the environmental variance component for measurement error, but we did not dissattenuate the test-retest correlation for measurement error.

Importantly, this analytic strategy is not without limitations. For example, Schmitt (1996) demonstrated that dissattenuation based on Cronbach's alpha can produce illogical estimates of true correlations because alpha overestimates measurement error under several circumstances. In the current context, this bias can have the effect of underestimating environmentality and overestimating environmental stability. We provide both measurement error corrected and uncorrected estimates. Corrected estimates should be interpreted as lower bounds for environmentality and upper bounds for environmental stability. Uncorrected estimates should be interpreted as upper bounds for environmentality and lower bounds for environmental stability.

### Time Lag

Roberts and DelVecchio's (2000) meta-analysis treated time lag between measurement occasions as a confound to be controlled. It is nearly a truism that traits are less stable over longer periods of time. Fraley and Roberts (2005) specifically emphasized this point by reanalyzing the data with variation in time interval, rather than age of the participants, as the primary independent variable predicting differential stability. They detected time-based decay that varied depending on the initial age of assessment. In order to avoid potential confounding of age and time-lag effects, we control for the linear effect of time lag in all analyses. However, the emphasis of the current meta-analysis is on life-span age-based trends. In order to confine the scope of this article, we do not report results of analyses that explore more nuanced nonlinear functions of time lag (importantly, results that control for nonlinear effects of time lag are nearly identical to those presented here).

### Trait Differences

One of the remarkable findings of Roberts and DelVecchio (2000) is that there are practically no differences between personality traits in terms of differential stability. In contrast, some traits, notably extraversion and neuroticism, have been regarded as having a more sure biological component than other traits (Eysenck, 1992). Using the Big Five framework (Digman, 1990), we examine whether the stability of genetic and environmental effects differs across traits.

### Self- Versus Other Report

Informant type is another possible moderator of trait stability. Roberts et al. (2008) hypothesized that increased personal awareness of one's identity is responsible for increasing differential stability of personality traits with age. As one encounters more experiences where choices have to be made that lean toward higher extraversion (e.g., going to a party) compared to lower extraversion (e.g., staying at home), one's identity becomes more strongly associated with these choices. This developmental model implies that self-reports of personality may be more likely to show evi-



dence of increasing stability with age. In spite of this hypothesis, previous work has found little evidence that self- and peer reports differ for behavioral genetic analyses (Riemann, Angleitner, & Strelau, 1997).

### Broad Versus Narrow Measurement

A final potential moderator is whether personality traits are measured at a broad level, such as the Big Five, or a more narrow level, such as the facet level. Ferguson (2010) included this distinction in his meta-analysis and found a relatively small difference in stability between broad personality traits ( $r = .76$ ) and narrow traits ( $r = .86$ ). Similarly, it may be the case that the stabilities of genetic and environmental influences on broad and narrow indices of personality differ. As we have been unable to identify a clear theoretical rationale for expecting such differences, we do not make specific empirical predictions about the direction of such differences and treat this question as exploratory. We do note however, that there is evidence that specific, facet-level personality factors may be influenced by a simpler, genetic architecture (e.g., Briley & Tucker-Drob, 2012; Terracciano, Sanna, et al., 2010) and that blended traits, such as integrity, have been put forward as better predictors of key life outcomes (e.g., Ones & Viswesvaran, 2001). Given that differences have been found between broad and narrow measures of personality in these other areas of inquiry, it is possible that differences might be found with respect to the stabilities of genetic and environmental effects on broad and narrow measures.

### Turning Points

Much emphasis has been placed on specific developmental transitions in personality maturation, the most famous being that personality stability is achieved at age 30 (McCrae & Costa, 1994; cf. Srivastava, John, Gosling, & Potter, 2003). However, there have also been accounts that continuous, progressive trends may better represent personality development. For example, Caspi and colleagues (2005) noted that “the level of stability increases in a relatively linear [read continuous] fashion through adolescence and young adulthood” (p. 467). Importantly, because differential stability is inherently a population-level statistic that indexes the ordering of individuals relative to one another, between-person heterogeneities in the timing and rate of life transition can result in a smoothed population-level age trend in differential stability in spite of discontinuities in individual-level developmental trajectories. For the current meta-analysis, we provide results of two analytic approaches: a linear spline approach, where turning points are selected on a priori grounds on the basis of typical transitions previously identified as meaningful in the literature, and a continuous parametric approach, which represents development more progressively. In the current context, the spline and continuous parametric approaches should be viewed as complementary descriptive accounts of the data, rather than competing models to be tested against one another.

We model turning points for each of the developmental trends at the ages of 3 years (temperament to traits),<sup>3</sup> 15 years (childhood to puberty), 30 years (adolescence to adult roles), and 60 years (transition to retirement). Personality research has typically classified measures that focus on infants and toddlers as temperament

and measures designed for adults as trait measures. However, this distinction may not be theoretically meaningful as both assessment strategies attempt to uncover systematic patterns of behavior across situations or time (McCrae et al., 2000). Empirically, temperament has been consistently linked with traitlike measures (Caspi et al., 2003; Measelle, John, Ablow, Cowan, & Cowan, 2005; Shiner, Masten, & Tellegen 2002). Other transitions such as early adolescence and associated pubertal changes (Arnett, 2000; McClintock & Herdt, 1996) and retirement and associated aging processes (Lucas & Donnellan, 2011; Roberts & DelVecchio, 2000) have been put forward. Each of these periods marks important social and biological transitions that may display different patterns of personality development. Although not often broken down specifically in this way, the FFT would be likely to view age 30 as an important turning point and the others to be relatively inconsequential. Conversely, SPMs would be more likely to endorse each turning point as important due to the shifts in environmental experiences.

### Goals of the Current Study

The current project sought to provide an exhaustive, quantitative review of longitudinal, behavioral genetic studies of personality development. This study is the first comprehensive empirical meta-analysis of this literature. The only previous synthesis of the literature was by Kandler (2012). In a short review of the longitudinal, behavioral genetic literature on extraversion and neuroticism, he found evidence that both genetic and environmental mechanisms will be necessary to explain personality development. However, the article mentioned some important limitations such as being a selective review of the literature, being a brief and largely descriptive and informal analysis of the data, and being limited to only two personality traits. The current project encompasses formal rigorous meta-analytic modeling of data obtained from an exhaustive aggregation of published studies on all personality traits. We examine effect sizes that fall into three classes: (a) the levels of heritability and environmentality of traits at one point in time, (b) the test–retest stability of phenotypic traits and of genetic and environmental effects, and (c) the contribution of genetic and environmental effects to test–retest stability. On the basis of our review above, the FFT would predict that genetic effects are large (i.e., high heritability), are stable (i.e., high  $rA$ ), and explain increases in phenotypic stability (i.e., age trends in genetic contribution parallel to age trends in phenotypic stability). According to this perspective, environmental effects are likely measurement error (i.e., substantial environmentality, but very low corrected environmentality), time-point-specific (i.e., low  $rE$ ), and unstructured in terms of the life course (i.e., minimal environmental contribution to age trends in phenotypic stability). SPMs, on the other hand, would predict that genetic influences on personality and increases in stability exist (i.e., nontrivial estimates of heritability,  $rA$ , and genetic contribution), but increases in phenotypic stability would be primarily mediated by increasingly stable environmental factors (i.e., increasing environmentality and  $rE$  with age).

<sup>3</sup> In the studies meta-analyzed for this article, age 3 years was the last age at which temperament was specifically measured in young children.



## Method

Following the Meta-Analysis Reporting Standards (Cooper, 2010), we report the search terms and databases used for identifying individual studies, inclusion and exclusion criteria, extraction and coding of effect sizes, and our statistical approaches for handling multiwave studies and modeling moderators of effect sizes (e.g., age).

### Data Aggregation

We performed an abstract search of PsycINFO for studies that included any combination of terms from three categories: genetics (*twin, genetic, adoption, adopted, adoptee*), methodology (*longitudinal, aging, stability*), and personality (*personality, temperament, trait*). This search produced 578 potential articles. Articles written in a language other than English or that sampled a clinical population were removed. To be included in the current study, the article had to provide information from which within- and across-time sibling group correlations could be derived (either raw or implied by a behavioral genetic model), compare siblings of similar ages (rather than parent-child correlations), and assess personality traits through self- or informant report formats. As is common with large, longitudinal studies, multiple articles were published using similar, updated data. When this was the case, we removed redundant articles and kept the publication with the most time points or the most measures. In order to test whether effect sizes differ for broad or narrow measurement of personality, we included studies that used the same data but reported the results at different levels of trait generality (e.g., Blonigen, Carlson, Hicks, Kreuger, & Iacono, 2008; Hopwood et al., 2011). Our final sample of studies included 24 longitudinal twin-sibling studies from 21 unique samples comprising 21,057 sibling pairs. The sibling pairs included 7,787 monozygotic reared-together twins, 12,951 dizygotic reared-together twins, 59 monozygotic reared-apart twins, 156 dizygotic reared-apart twins, 60 half-siblings reared together, and 44 biologically unrelated siblings reared together. Table 1 presents the citation, data set, age ranges, measures, and sample size of each article. Raw or model implied group within- and across-time correlations were extracted from each article for each pair of time points and repeated measures.<sup>4</sup>

We used Mplus statistical software (Muthén & Muthén, 1998–2010) to fit a longitudinal correlated factors model (see Figure 2, described above) for each outcome from each study individually using the summary data option that allowed us to estimate the parameters based on the extracted group correlation matrices. To obtain accurate standard errors, we input the sample size that each group contained in terms of complete sibling pairs that were present at least at two time points. In practice, this approach proved difficult as the reporting of sample size varied across study. The current analysis uses the sibling pair as the unit of analysis, but sample size was often reported in terms of individuals. Additionally, some studies continued to recruit new participants as the study progressed, rendering it difficult to interpret how many twin pairs had complete longitudinal data. Because of this uncertainty, we chose to take a conservative approach to estimating sample size. When sample size was only reported in terms of individuals, we took half the number to represent the number of pairs. When available, we used the number of complete pairs that had full data for at least two waves of data. When unavailable, we used reported

attrition statistics to calculate the number of pairs that completed at least two assessments. Sample attrition or continued sampling is unlikely to exert a large influence on the current results as the majority of studies ( $K = 15$ ) only reported information for two waves, and the maximum number of waves was four. Estimating effect sizes from raw data would likely have resulted in smaller standard errors, as access to individual level data allows for powerful techniques that can handle missing data (e.g., full-information maximum-likelihood estimation). Therefore, the reported results may be considered lower bound estimates of precision.

Having completed this process, we compiled each parameter depicted in Figure 2 with the associated standard error and calculated the phenotypic stability and the genetic and environmental contribution to stability and associated standard errors. We encountered a few instances where the correlated factors model applied to summary data converged to an out-of-bounds estimate. Correlations are bounded by  $-1$  and  $1$ , but some estimates of genetic stability and corrected environmental stability exceeded the upper bound. This result occurred for a small number of effect sizes. For genetic stability, nine estimates were greater than  $1$  (mean estimate =  $1.09$ , range =  $1.01$ – $1.28$ ). For corrected environmental stability, 31 estimates were greater than  $1$  (mean estimate =  $1.39$ , range =  $1.02$ – $2.61$ ). These estimates tended to be imprecise (mean  $SE = 0.20$ ). As these estimates tended to carry comparatively small weight in the analyses, we chose to leave them in the data set as estimated. The patterns of results are unchanged if these estimates are removed or censored to the closest reasonable estimate.

We formed two data sets: one arranged in a cross-sequential manner with heritability and environmentality at each time point and another with the phenotypic stability and genetic and environmental correlations and contributions to phenotypic stability associated with each *pair* of time points. This approach resulted in  $330 \times 3$  (heritability, environmentality, and corrected environmentality) effect sizes for the cross-sequential data set and  $251 \times 6$  (phenotypic stability, genetic stability, environmental stability, corrected environmental stability, genetic contribution to stability, and environmental contribution to stability) effect sizes in the longitudinal data set. In the cross-sequential data set, each study contributed an average of 28.66 sets of effect sizes ( $SD = 17.74$ , range =  $2$ – $60$ ). In the longitudinal data set, each study contributed an average of 26.26 sets of effect sizes ( $SD = 20.17$ , range =  $1$ – $60$ ). These outcomes were associated with information about age in the first data set and age at the initial time point and the time interval between measurements in the second data set. Additional variables included in the data set are described below.

### Study/Variable Characteristics

**Error correction.** For most articles ( $K = 16$ ), we were able to obtain calculated reliability estimates (Cronbach's alpha) within the sample at each time point and for each measure, which is

<sup>4</sup> We determined that latent growth curve models were too restrictive to accurately extract group correlations for pairs of time points, but we included common longitudinal behavioral genetic models such as the Cholesky decomposition, the correlated factors model, and the simplex model, as well as variations of these.

Table 1  
Description of Included Studies

Citation	Data set	Variables examined	Number of waves	Age range (years)	Longitudinal sample size (pairs)
Blonigen, Carlson, Hicks, Kreuger, & Iacono (2008) <sup>a</sup>	Minnesota Twin Family Study	Absorption, achievement, aggression, alienation, control, harm avoidance, social closeness, social potency, stress reactivity, traditionalism, well-being	2	17.00–24.00	360 MZT, 188 DZT
Blonigen, Hicks, Kreuger, Patrick, & Iacono (2006) <sup>a</sup>	Minnesota Twin Family Study	Fearless dominance, impulsive antisociality	2	17.00–24.00	360 MZT, 188 DZT
Bratko & Butkovic (2007) <sup>b</sup>	Zagreb Community Sample	Extraversion, neuroticism, psychoticism, Lie Scale	2	17.00–21.00	75 MZT, 85 DZT
De Fruyt et al. (2006) <sup>c</sup>	Small-scale twin family study	Emotional instability, extraversion, imagination, benevolence, conscientiousness	2	8.65–11.65	79 MZT, 124 DZT
Ericson, Tuvblad, Raine, Young-Wolff, & Baker (2011) <sup>†</sup>	Southern California Twin Project	Cognitive-perceptual, disorganization, interpersonal-affective	2	11.89–14.69	205 MZT, 293 DZT
Forsman, Lichtenstein, Andershed, & Larsson (2008) <sup>b</sup>	Twin Study of Child and Adolescent Development	Grandiose/manipulative, callous/unemotional, impulsive/irresponsible	2	16.00–19.00	307 MZT, 473 DZT
Gagne & Goldsmith (2011) <sup>c</sup>	Wisconsin Community Sample	Anger	2	1.00–3.00	130 MZT, 237 DZT
Ganiban, Saudino, Ulbricht, Neiderhiser, & Reiss (2008) <sup>c</sup>	Nonshared Environment and Adolescent Development	Activity, emotionality, shyness, sociability	2	12.80–15.45	63 MZT, 228 DZT, 60 HST, 44 URT
Gillespie, Evans, Wright, & Martin (2004) <sup>b</sup>	Brisbane Community Sample	Extraversion, neuroticism, psychoticism	3	12.00–16.00	216 MZT, 192 DZT
Hopwood et al. (2011) <sup>‡</sup>	Minnesota Twin Family Study	Constraint, negative emotionality, positive emotionality	3	17.00–29.00	349 MZT, 183 DZT
Johnson, McGue, & Kreuger (2005) <sup>b</sup>	Minnesota Twin Study of Adult Development and Aging	Absorption, achievement, aggression, alienation, control, harm avoidance, social closeness, social potency, stress reactivity, traditionalism, well-being, constraint, negative emotionality, positive emotionality	2	59.40–64.40	384 MZT, 274 DZT
Kandler et al. (2010) <sup>d</sup>	Bielefeld Longitudinal Study of Adult Twins (cross-sequential design)	Extraversion, agreeableness, conscientiousness, neuroticism, openness to experience	3	22.70–35.20 (young cohort) 41.20–55.00 (middle cohort) 17.20–29.50	156 MZT, 95 DZT (young cohort) 140 MZT, 88 DZT (middle cohort) 650 MZT, 650 DZT
Kupper, Boomsma, de Geus, Denollet, & Willemsen (2011) <sup>b</sup>	Netherlands Twin Registry	Negative affectivity, social inhibition	3		
McGue, Bacon, & Lykken (1993) <sup>b</sup>	Minnesota Community Sample	Absorption, achievement, aggression, alienation, control, harm avoidance, social closeness, social potency, stress reactivity, traditionalism, well-being, constraint, negative emotionality, positive emotionality	2	19.80–29.60	79 MZT, 48 DZT

Table 1 (continued)

Citation	Data set	Variables examined	Number of waves	Age range (years)	Longitudinal sample size (pairs)
Niv, Tuvblad, Raine, Wang, & Baker (2011) <sup>b</sup>	Southern California Twin Project	Inattention-impulsivity, motor impulsivity, nonplanning	2	11.89–14.69	205 MZT, 293 DZT
Pedersen & Reynolds (1998) <sup>b</sup>	Swedish Adoption/Twin Study of Aging	Extraversion, neuroticism, openness to experience	4	60.15–69.15	96 MZT, 123 DZT, 59 MZA, 156 DZA
Read, Vogler, Pedersen, & Johansson (2006) <sup>b</sup>	Octogenarian Twin Study	Extraversion, neuroticism	3	82.30–86.30	149 MZT, 202 DZT
Rietveld, Hudziak, Bartels, van Beijsterveldt, & Boomsma (2004) <sup>c</sup>	Netherlands Twin Registry	Attention problems	4	3.00–12.00	1,891 MZT, 3,310 DZT
Saudino (2012) <sup>c</sup>	Boston University Twin Project	Activity level	2	2.00–3.00	141 MZT, 163 DZT
Saudino & Cherny (2001) <sup>a, c</sup>	MacArthur Longitudinal Twin Study	Activity, affect-extraversion, behavioral inhibition, shyness, task orientation	4	1.17–3 .00	118 MZT, 106 DZT
Spengler, Gottschling, & Spinath (2012) <sup>b</sup>	Cognitive Ability and Self-Perceived Motivation of School Achievement	Extraversion, agreeableness, conscientiousness, neuroticism, openness to experience	2	9.09–13.10	67 MZT, 135 DZT
Takahashi et al. (2007) <sup>b</sup>	Keio Twin Project	Behavioral Inhibition Scale, Behavioral Activation Scale	2	23.27–25.52	85 MZT, 32 DZT
Viken, Rose, Kaprio, & Koskenvuo (1994) <sup>b</sup>	Finish Twin Cohort (cross-sequential design)	Extraversion, neuroticism	2 (multiple cohorts)	20.50–26.50 26.50–32.50 32.50–38.50 38.50–44.50 44.50–50.50 50.50–56.50	801 MZT, 1,572 DZT 580 MZT, 1,180 DZT 405 MZT, 798 DZT 288 MZT, 618 DZT 216 MZT, 498 DZT 171 MZT, 339 DZT 345 MZT, 895 DZT
Zavos, Gregory, & Eley (2012) <sup>b</sup>	Genesis 12-19 Study	Anxiety sensitivity	3	15.00–17.00	

Note. Citations marked with the same subscript letter use the same data. MZT = monozygotic twins reared together; DZT = dizygotic twins reared together; MZA = monozygotic twins reared apart; DZA = dizygotic twins reared apart; HST = half-siblings reared together; URT = unrelated siblings reared together.

<sup>a</sup> Data on affect-extraversion were not available for Wave 4, and thus, only data on ages 1.17 to 2 years were analyzed. <sup>b</sup> Indicates self-report format. <sup>c</sup> Indicates informant report format. <sup>d</sup> Indicates both self-report and informant report formats.



necessary to accurately correct the environmental component for measurement error. Four of these articles cited other work using the same sample and measures from which we obtained the estimates. However, some articles reported reliability as an average or range across waves ( $K = 4$ ) or across traits ( $K = 1$ ), only reported initial reliability ( $K = 1$ ), or only reported estimates of reliability from a manual ( $K = 2$ ). Although not ideal, we extracted the maximum amount of information possible and associated it with the personality information. We took the midpoint of a range as the best estimate of reliability or the average when given. Initial reliability estimates were carried forward in time to apply to later time points in which reliability information was unavailable. The average reliability for all time points and measures was .78 ( $SD = .10$ , range = .30–.94).

There was one other curious case that highlights a potential limitation of this approach. [Saudino \(2012\)](#) reported a reliability coefficient of .78 for parental reports of activity level and an intraclass monozygotic twin correlation of .82, rendering a correction for measurement error unidentified (i.e., negative environmental variance). This effect size for parent report of activity level was dropped when correcting for measurement error.

**Trait differences.** A diverse array of personality instruments was used in the identified studies. These included the Eysenck Personality Questionnaire ([Eysenck & Eysenck, 1975](#)), the Hierarchical Personality Inventory for Children (HiPIC; [Mervielde & De Fruyt, 1999](#)), the Youth Psychopathic Traits Inventory ([Andershed, Kerr, Stattin, & Levander, 2002](#)), the EAS Temperament Survey ([Buss & Plomin, 1984](#)), the Multidimensional Personality Questionnaire (MPQ; [Tellegen & Waller, 2008](#)), the NEO Five-Factor Inventory ([Costa & McCrae, 1989](#)), the NEO Personality Inventory—Revised ([Costa & McCrae, 1992](#)), the Behavioral Inhibition Scale/Behavioral Activation Scale (BIS/BAS; [Carver & White, 1994](#)), the Child Anxiety Sensitivity Index ([Silverman, Fleisig, Rabian, & Peterson, 1991](#)), the Infant Behavior Record ([Bayley, 1969](#)), the Infant Behavior Questionnaire ([Rothbart, 1981](#)), the Children's Behavior Questionnaire ([Rothbart, Ahadi, Hershey, & Fisher, 2001](#)), the Child Behavior Checklist ([Achenbach, 1991](#)), the Toddler Behavior Assessment Questionnaire ([Goldsmith, 1996](#)), the Barratt Impulsiveness Scale ([Patton, Stanford, & Barratt, 1995](#)), the Type D scale ([Kupper & Denollet, 2007](#)), and the Schizotypal Personality Questionnaire—Child ([Raine & Baker, 2003](#)).

We used the Big Five taxonomy to organize these various scales. Each effect size was coded as indicative of extraversion, agreeableness, conscientiousness, neuroticism, or openness to experience based on conceptual and empirical links between different measures (e.g., [Church, 1994](#); [Donnellan, Conger, & Burzette, 2005](#); [John, Naumann, & Soto, 2008](#), p. 115). For example, we coded the positive emotionality, negative emotionality, constraint, and absorption subscales of the MPQ as extraversion, neuroticism, conscientiousness, and openness, respectively. For the HiPIC, we coded the emotional instability, extraversion, imagination, benevolence, and conscientiousness subscales as neuroticism, extraversion, openness, agreeableness, and conscientiousness, respectively. For the EAS Temperament Survey, we coded the negative emotionality subscale as neuroticism and the activity level, sociability, and shyness subscales as extraversion. For the BIS/BAS, we coded behavioral inhibition as neuroticism and behavioral activation as extraversion. Generally, we coded subscales related to the

ability to focus (e.g., task orientation, impulsivity, inattention) as conscientiousness,<sup>5</sup> and we coded subscales of normal-range measures related to psychopathy (e.g., cognitive-perceptual deficits, Type D) or emotional instability (e.g., anger, anxiety sensitivity) as neuroticism. We used a consensus method to label the constructs and reached agreement. We used effects coding to create four variables with extraversion as the reference trait (coded as  $-1$ ) to contrast with the specified trait (coded as  $1$ ) and the nonspecified traits (coded as  $0$ ). Extraversion was the most studied trait ( $k = 98$  across all time points),<sup>6</sup> followed by neuroticism ( $k = 94$ ), conscientiousness ( $k = 67$ ), agreeableness ( $k = 43$ ), and openness to experience ( $k = 30$ ).

**Self- versus other report.** Effect sizes based on data from observer, parent, informant, or generally someone other than the target were coded as  $-.5$  for an effects coded variable, and effect sizes based on self-report were coded as  $.5$ . The majority of assessment strategies were self-report ( $k = 253$ ), with a sizable minority using other report ( $k = 77$ ).

**Broad versus narrow.** We classified effect sizes based on traits at the level of the Big Five or broader as broad measures and facet or more specific constructs as narrow measures. For example, we treated the superfactors of the MPQ (e.g., constraint, negative emotionality, positive emotionality, absorption; [Tellegen & Waller, 2008](#)) as broad measures and the more specific scales (e.g., achievement, control, harm avoidance, etc.) as narrow measures. For studies that focused on a specific trait not in reference to a general taxonomy (e.g., anxiety sensitivity, impulsivity, shyness, etc.), we coded these as narrow measures. We created an effects coded variable with narrow measures coded as  $-.5$  and broad measures as  $.5$ . Slightly more measures were classified as narrow ( $k = 170$ ) than broad ( $k = 160$ ).

## Analytic Approach

Our primary goal was to test a series of alternative models that specified different life-span trends for the genetic and environmental influences on personality development. To accomplish this goal, we fit random-effects, meta-analytic regression models using the general framework laid out by [Cheung \(2008\)](#). Random-effects models are considered the most conservative and therefore preferred modeling strategy for meta-analytic studies. Random-effects models estimate the error associated with an effect size as well as variation in the across-study true effect size ([Hedges & Vevea, 1998](#)). To examine the possibility of continuous growth rather than growth patterns that apply to specific age ranges, we examined exponential models. Exponential models are continuous parametric models, but because they are parametrically nonlinear, it is computationally unfeasible to fit them as random-effects models with currently available software. These models were fit as fixed-effects models. Therefore, we offer these models as a comparison to evaluate whether the preferred linear models approximate a

<sup>5</sup> Because the behavioral genetic models used in the current study are primarily concerned with sibling correlations, not means, our results are unaffected by giving a similar label to forward- and reverse-coded scales.

<sup>6</sup> All  $k$ s reported for the moderators refer to sets of effect sizes associated with specific time points rather than pairs of measurement.

continuous trend or display important deviations from the continuous trend.<sup>7</sup>

To obtain accurate standard errors of our regression estimates, we used a weighting procedure. All variables were weighted by the inverse of the sampling variance of the dependent variable and the inverse of the number of effect sizes included per study. The cluster option of Mplus was applied to correct for nonindependence of effect sizes drawn from the same sample. By using these procedures, our estimates are not biased by including multiple effect sizes from the same sample.

We constructed connected-linear and continuous exponential models of (a) age trends in the level of heritability, environmentality, and corrected environmentality across the life span; (b) age trends in stability at the phenotypic, genetic, environmental, and corrected environmental level across the life span; and (c) age trends in the contribution of genetic and environmental effects to phenotypic stability across the life span. For the first set of analyses (in which the outcomes are specific to individual time points), each outcome was predicted by the associated age. For the second and third sets of analyses (in which the outcomes are derived from longitudinal pairs of time points), the outcome was predicted by age at baseline controlling for time interval between measurements. In all models, we did not adjust the coding of participant age, and therefore, the intercept of the model applies to age 0. As the earliest effect size occurs at age 1, this approach requires a very minor extrapolation. Additionally, we explicitly vary participant age in our interpretation and figures. In all models based on longitudinal data, we control for the linear effect of time lag, which was centered at the across sample average of 5.563 years. Thus, all interpretation and graphical presentation of the age-related effects assume a time lag of roughly 5.6 years.

We were interested in the possibility of there being different age trends during different developmental periods. To examine the functional form of the developmental trends in greater detail, we fit a series of connected-linear spline models for each outcome. We selected the ages of 3, 15, 30, and 60 years as turning points. As discussed earlier, these ages have been highlighted in previous research and represent the transition from temperament into childhood, the transition from childhood into emerging adulthood, the transition into fully mature roles such as work and marriage, and the transition out of work. In total, we fit 17 linear models for each outcome: the mean effect size, a linear model with a single slope, and every logical combination of the break points from a model containing two slopes to a model containing five slopes. These variations allow for a full examination of whether the age trends in the outcomes differ depending on developmental period in a manner that our continuous exponential model would not be able to detect. To compare the results of the linear spline models with a continuous function, we fit an exponential model to the aggregated data for each outcome. This model had the form of

$$\hat{y}_i = b_0 - b_1 e^{b_2(\text{age}_i)},$$

where  $\hat{y}_i$  represents a predicted effect size,  $b_0$  represents the horizontal asymptote,  $b_1$  represents a scaling factor, and  $b_2$  represents the growth rate.

We report the full connected-linear model that is the most unrestricted account of the data (i.e., has five slopes), but this model is likely unnecessarily complicated. To compare models, we take into account the fit statistics of loglikelihood (LL), the Akaike

information criterion (AIC), and the Bayesian information criterion (BIC). In each case, statistics that are closer to zero indicate better model fit. LL represents the absolute amount of misfit between the model and the actual data, whereas AIC and BIC take into account both misfit and model parsimony, with BIC weighting parsimony to a greater extent. Preferred models are selected on the basis of AIC and BIC and interpreted for each outcome. In cases where AIC and BIC differed, we report and examine both models but place emphasis on the BIC preferred model as BIC represents a more parsimonious account of the data (Kline, 2011; Vrieze, 2012). We take this approach in order to balance the strengths and weaknesses of connected-linear models. A strength of connected-linear models is that age-specific developmental trends can be identified, but an associated weakness is that simple noise in the data can appear as meaningful developmental differences. Connected-linear models are subject to overparameterization that leaves the developmental trend jagged and somewhat difficult to interpret. Therefore, interpretation of the more parsimonious BIC preferred model can avoid some of these issues. The most straightforward way to avoid these issues is to examine the scatterplots of the data. We plot each outcome based on the associated age and weight the data points based on the weighting scheme described earlier. The preferred connected-linear and continuous exponential models are overlaid on the data to aid interpretation.

We also sought to determine if effect sizes differed based on the identified potential moderators. Therefore, in a follow-up set of analyses, we included the set of effects coded variables for the Big Five trait categories, self- versus other report, and broad versus narrow measurement as predictors in separate models. Finally, we report two sensitivity analyses, the first to examine the effect of including only self-report effect sizes and the second to probe for publication bias.

## Results

### Descriptive Statistics

Table 2 presents descriptive statistics of the studies weighted by the number of sibling pairs per study. Average age at baseline and follow-up assessment was in the 20s with sizable standard deviations. The studies covered nearly the entire range of the life span, beginning at 1 year of age and ending with 86 years. The average time interval between assessments was 5.43 years. A little more than half of the participants were female. The racial/ethnic composition of the samples was not well reported in the articles. On the

<sup>7</sup> We chose an exponential model as our preferred nonlinear model for theoretical and practical reasons. Other possible nonlinear models include quadratic and dual exponential models. Applied to the expected trend of large increases in stability at young ages followed by a plateau in adulthood, an exponential model resembles this trend the most. Quadratic and dual exponential models would allow for an upward trajectory in early life with an associated decline in old age. This trend has been found for personality stability in old age (Lucas & Donnellan, 2011). However, our coverage of this region was very sparse, rendering it difficult for our continuous models to pick up on this slight trend. Practically, quadratic models applied to the data rarely allowed for convergence, and dual exponential models tended to fit worse than more parsimonious exponential models. For these reasons, we only report the results from the continuous exponential growth models.

Table 2  
*Characteristics of the Studies Meta-Analyzed*

Variable	<i>M</i>	<i>SD</i>	Range
Age at baseline (years)	23.02	19.99	1.00–84.30
Interval (years)	5.43	2.91	0.33–13.80
Age at followup (years)	28.45	20.53	1.67–86.30
Percent female	54.12	9.46	0.00–100.00
Percent non-White	24.09	31.34	2.00–73.00

*Note.* Estimates weighted by sample size. Percent non-White may be misleading as many studies did not report racial composition of the sample and were therefore coded as missing. However, given the populations sampled from (e.g., Minnesota, Northern Europe), it is likely that the large majority of the participants were White.

basis of the location and populations sampled (e.g., Minnesota, Northern Europe), we can infer that the twin studies tended to be predominantly White. The estimate given in Table 2 is likely to be an overestimate of the minority representation in the samples because the Southern California Twin Project is a large outlier in terms of racial composition with over 70% minority participants (Niv, Tuvblad, Raine, Wang, & Baker, 2012). This project contrasts with the large-scale European studies in Finland (Viken, Rose, Kaprio, & Koskenvuo, 1994) and the Netherlands (Rietveld, Hudziak, Bartels, van Beijsterveldt, & Boomsma, 2004) that contain vastly larger sample sizes but do not report information about racial or ethnic background. Information about age was complete, and information regarding gender composition was relatively complete.

### Age Trends in the Level of Heritability and Environmentality

Table 3 presents the results for the random-effects, meta-analytic connected-linear spline regression models predicting levels of heritability, environmentality, and corrected environmentality by age. Table 4 presents the results of the fixed-effects, exponential models for the outcomes. Figure 3 presents the raw data and best fitting models graphically. In Figure 3, the scatterplot of the data is displayed with the trend lines superimposed. Each effect size is represented as a dot at the point estimate surrounded by a circle with a diameter that has been scaled relative to the weight that the data point carried in the analyses. To aid visualization, we performed transformations on the weighting variable (e.g., dividing by a constant, taking the square root to minimize outliers), and an upper limit was placed on the size of each circle. Larger circles are indicative of data points that were weighted more strongly.

The mean effect-size model indicates that roughly half of the variance in personality on average can be attributed to genes (48%) and half to the environment (52%), which is consistent with previous research (Bouchard & Loehlin, 2001). However, there are significant age trends across the life span. The full connected-linear model for heritability indicates substantial genetic influences very early in life, followed by a drop-off in early childhood and plateau in middle age. Environmentality displays the opposite trend. However, not all of these slopes were necessary to account for the age trends present in the data. To represent the reduced models in a consistent tabular form, we have placed arrows in the

Table 3  
*Age Trends for Heritability, Environmentality, and Corrected Environmentality*

Regression model	$b_0$	$b_{0-.3}$	$b_{3-.15}$	$b_{15-.30}$	$b_{30-.60}$	$b_{60-.90}$	$\tau$	LL	AIC	BIC
<b>Heritability</b>										
Mean effect size	.480 (.020)****	—	—	—	—	—	.092 (.016)****	–390.508	785.016	792.614
Full model	.547 (.129)****	.041 (.047)	–.015 (.005)**	–.004 (.003)	—	—	.061 (.012)****	–372.163	758.326	784.920
Preferred model (AIC)	.541 (.131)****	.044 (.049)	–.017 (.005)****	–.001 (.001)*	←	–.002 (.001)*	.063 (.013)****	–372.727	755.455	774.450
Preferred model (BIC)	.676 (.068)****	→	–.013 (.005)**	–.001 (.001)*	←	←	.065 (.013)****	–373.908	755.816	771.012
<b>Environmental</b>										
Mean effect size	.515 (.020)****	—	—	—	—	—	.091 (.016)****	–395.508	795.016	802.614
Full model	.453 (.128)****	–.041 (.047)	.014 (.006)**	.005 (.003)	—	—	.061 (.012)****	–376.736	767.473	794.067
Preferred model (AIC)	.460 (.131)****	–.045 (.048)	.016 (.005)**	.002 (.001)*	←	←	.063 (.012)****	–377.461	764.921	783.917
Preferred model (BIC)	.324 (.067)****	→	.013 (.005)*	.002 (.001)*	←	←	.065 (.013)****	–378.660	765.320	780.516
<b>Corrected environmental</b>										
Mean effect size	.311 (.023)****	—	—	—	—	—	.083 (.017)****	–388.250	780.499	788.085
Full model	.307 (.194)	–.042 (.077)	.009 (.006)	.006 (.004)	←	–.002 (.001)	.066 (.013)****	–377.545	769.090	795.641
Preferred model (AIC and BIC)	.187 (.053)****	→	→	.006 (.002)**	–.001 (.001)	←	.067 (.013)****	–378.445	764.981	780.063

*Note.* Dashes indicate that the parameter was not estimated.  $b_0$  represents the intercept,  $b_{0-.3}$  represents the coefficient for the first linear segment,  $b_{3-.15}$  represents the following segment, and so on until  $b_{60-.90}$ . For simplified models, arrows are used to indicate the parameter estimate that applies during the given developmental period.  $\tau$  represents the random-effect standard deviation in the between-study effect size. AIC = Akaike information criterion; BIC = Bayesian information criterion; LL = loglikelihood. \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ . \*\*\*\*  $p < .0001$ .



Table 4  
Results of Continuous Exponential Models for Each Developmental Outcome

Outcome	b <sub>0</sub>	b <sub>1</sub>	b <sub>2</sub>	b <sub>time lag</sub>	LL	AIC	BIC
Heritability	.336 (.074)***	-.405 (.068)***	-.037 (.024)	—	-418.390	2,916.778	2,928.175
Environmentality	.665 (.074)***	.406 (.068)***	-.037 (.023)	—	-421.816	2,925.297	2,936.694
Corrected environmentality	.451 (.074)***	.388 (.089)***	-.033 (.020)	—	-1,493.071	2,997.071	3,016.067
Phenotypic stability	.706 (.048)***	.349 (.075)***	-.055 (.022)*	-.006 (.006)	-414.479	4,944.914	4,959.016
Genetic stability	.997 (.016)***	.506 (.040)***	-.070 (.015)***	-.011 (.004)**	-290.729	4,405.366	4,419.468
Environmental stability	.539 (.779)	.403 (.594)	-.080 (.808)	-.025 (.012)*	-393.141	4,421.708	4,435.810
Corrected environmental stability	1.014 (.168)***	.767 (.132)***	-.029 (.021)	-.017 (.010)	-357.949	3,845.433	3,859.519
Genetic contribution to stability	.382 (.029)***	.356 (.140)*	-.860 (.210)***	-.003 (.007)	-290.733	4,321.622	4,335.707
Environmental contribution to stability	.330 (.044)***	.306 (.039)***	-.045 (.018)*	-.007 (.003)*	-281.284	4,798.369	4,812.455

Note. AIC = Akaike information criterion; BIC = Bayesian information criterion; LL = loglikelihood.  
\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

table representing whether an earlier or later parameter estimate is applied to the age range. Adjacent parameter estimates represent alterations of the slope of the developmental trend. For example, the BIC preferred model of heritability indicated that only two slopes were required to properly describe the trend. One slope is

for the ages of 0 to 15, and therefore, an arrow has been placed in the b<sub>0-3</sub> column pointing toward the parameter estimate in the b<sub>3-15</sub> column, indicating that this slope applies in the age range of 0 to 3. Similarly, arrows pointing toward the earlier time point are presented for Columns b<sub>30-60</sub> and b<sub>60-90</sub>, indicating that the b<sub>15-30</sub>

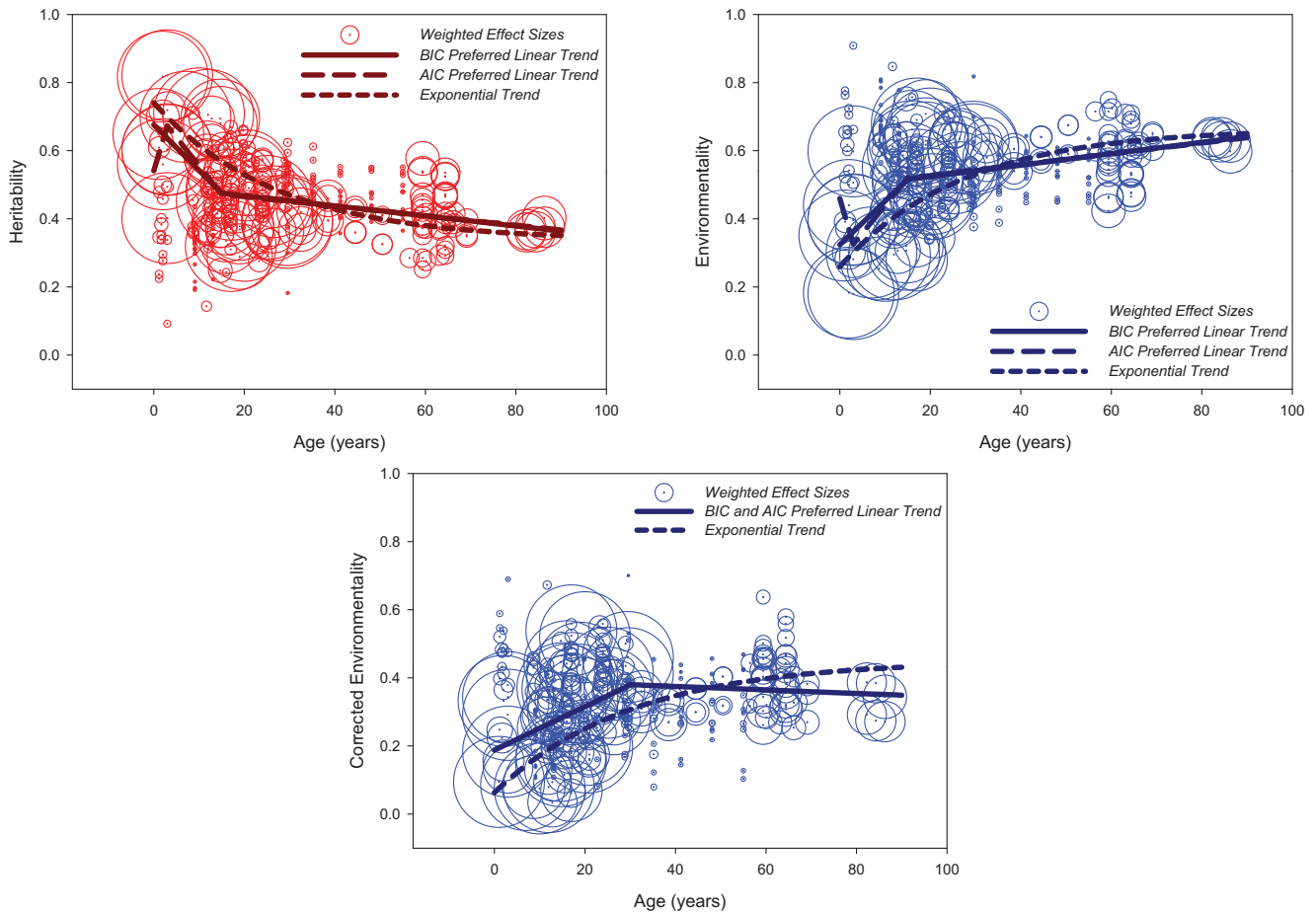


Figure 3. Age trends in heritability, environmentality, and measurement error corrected environmentality. Circles surrounding data points are scaled by the weighting variable (described in the Analytic Approach section in the text) such that larger circles carried more weight in the analysis. AIC = Akaike information criterion; BIC = Bayesian information criterion.

slope applies during this area of the life span. Because parameter estimates are adjacent in the  $b_{3-15}$  and  $b_{15-30}$  columns, these can be interpreted as turning points in the spline models.

In evaluating which reduced model for heritability and environmentality possessed the proper balance of parsimony and accuracy, the AIC and BIC differed in that the AIC indicated that a break point was needed at age 3. The BIC preferred model indicates that two linear slopes, one before age 15 and one after age 15, are required to accurately reflect the data, and we focus on this model. Near birth, genetic influences account for nearly 70% of variation in personality, and environmentality only accounts for 30% of variation. However, genetic effects decrease ( $b_{0-15} = -.013, p < .01$ ) in magnitude while environmental effects increase ( $b_{0-15} = .013, p < .05$ ) across early childhood. Following age 15, relatively small but significant changes occur in genetic ( $b_{15-30} = -.001, p < .05$ ) and environmental ( $b_{15-30} = .002, p < .05$ ) effects. This result is somewhat different when the environment is corrected for measurement unreliability. For example, the average effect size indicates that 31% of variation in personality associated with true environmental effects rather than 52%. The model that best reflects trends in corrected environmentality has different slopes before and after age 30. Prior to age 30, true environmentality increases ( $b_{0-30} = .006, p < .01$ ) from accounting for roughly 20% of the variance in early childhood and remains constant after age 30 ( $b_{30-90} = -.001, p = .63$ ).

The results of the continuous exponential models are largely similar. Heritability begins high in early life and declines to approach a lower asymptote of .34, whereas environmentality begins low and increases to approach an upper asymptote of .67 (.45 when corrected for measurement error). The asymptotic levels are approached rather gradually. At age 15, the models imply levels of heritability, environmentality, and corrected environmentality of .59, .43, and .21, respectively. By age 30, the genetic and environmental influences on personality are still relatively far from their asymptotic levels, with the estimate of heritability declining to .47, environmentality rising to .53, and corrected environmentality rising to .31.

One can visualize the similarities and differences between the age trends indicated by the different models by inspecting the plotted trend lines in Figure 3. Whereas the linear models predict a relatively constant level of genetic and environmental effects following age 15, the exponential model displays steady change until later in the life span. However, the different models are highly similar, they overlap for much of the life span, and each visually reflects the weighted data accurately. In comparing the different AIC and BIC preferred models, it is apparent that they nearly completely overlap. The major difference concerns the age trend for very young children, but beyond this discrepancy, the models indicate the same developmental trajectory. Again, very little difference is observed between the connected-linear trend and the continuous exponential trend apart from slight deviations in adolescence. This result may indicate that the connected-linear models may have an advantage over the exponential models as the scatterplot does seem to indicate more rapid change during this time period. On the whole, however, similar developmental implications can be derived from both analytical strategies. Heritability tends to be high and accounts for the large majority of variance in infancy. Heritability declines substantially in early childhood and somewhat more slowly thereafter, resulting in roughly 40% of

variance in personality attributable to genetic influences throughout adulthood. Environmentality displays the opposite pattern. Corrected for measurement error, environmentality shows a similar upward trajectory, with a peak of accounting for over 40% of variance in personality.

### Age Trends in Phenotypic, Genetic, and Environmental Stability

Table 5 presents the connected-linear model results for genetic and environmental stability as predicted by age at baseline and time interval between measurements. Table 4 presents the results of the continuous exponential model. Figure 4 presents the preferred trend lines and meta-analytic data graphically.

We begin by reviewing the results from the linear spline models. On average, test-retest stability was relatively high ( $b_0 = .554, p < .001$ ). Genetic stability ( $b_0 = .844, p < .001$ ) was substantially greater than both environmental ( $b_0 = .397, p < .001$ ) and corrected environmental stability ( $b_0 = .652, p < .001$ ). However, stability differed substantially over the life span. For phenotypic stability, the full model indicated that in infancy, differential stability was not different from zero ( $b_0 = .093, p = .47$ ). Stability rose sharply over the first 3 years of life ( $b_{0-3} = .126, p < .05$ ), plateaued during childhood ( $b_{3-15} = .001, p = .88$ ), and then resumed increasing substantially until age 30 ( $b_{15-30} = .016, p < .001$ ). Following age 30, no slope was significantly different from zero, but there was a small trend toward decreasing stability in old age. The model comparison indicated that the slopes for ages 30–60 and 60–90 could be constrained to be equal, but every other slope was needed.

A different story emerged for genetic and environmental stability. Genetic stability only required two slopes. This model indicated that genetic stability was high in infancy ( $b_0 = .553, p < .001$ ) and increased linearly until age 30 ( $b_{0-30} = .016, p < .001$ ). At this age, genetic stability reached unity and remained nearly perfectly stable across the remainder of the life span. The preferred model for environmental stability was slightly more complex. In infancy, environmental stability rises quickly ( $b_{0-3} = .176, p < .001$ ) and then continues to rise at a slow, steady rate over the majority of the life span ( $b_{3-60} = .006, p < .001$ ). Following age 60, environmental stability appears to decrease. However, this decrease appears to be attributable to measurement error. The age trends for measurement error corrected environmental stability display a somewhat jagged trend (i.e., rises in infancy, declines in childhood, and then rises in adolescence), but following age 30, environmental stability remains constant ( $b_{30-90} = .002, p = .22$ ).

The continuous exponential models are largely consistent with the life-span trends as indicated by the connected-linear models. Each model indicated growth across the life span, with an upper asymptote of .71 for phenotypic stability, 1.00 for genetic stability, .54 for environmental stability, and 1.01 for corrected environmental stability. Growth in phenotypic, genetic, and environmental stability was largely concentrated at early ages. The models indicate that phenotypic, genetic, and environmental stabilities are .55, .82, and .42, respectively, by age 15. At 30 years of life, the expected phenotypic, genetic, and environmental stabilities will have nearly approached their asymptotic levels and are predicted to be .64, .94, and .50, respectively. Thus, consistent levels of stability are predicted across adulthood and old age. Corrected

**Table 5**  
**Age Trends for Phenotypic, Genetic, Environmental, and Corrected Environmental Stability**

Regression model	$b_0$	$b_{0-3}$	$b_{3-15}$	$b_{15-30}$	$b_{30-60}$	$b_{60-90}$	$b_{\text{time lag}}$	$\tau$	LL	AIC	BIC
<b>Phenotypic stability</b>											
Mean effect size	.554 (.036) <sup>***</sup>	—	—	—	—	—	—	.100 (.016) <sup>***</sup>	-403.240	810.479	817.530
Full model	.093 (.123)	.126 (.051) <sup>*</sup>	.001 (.004)	.016 (.003) <sup>***</sup>	.000 (.002)	-.003 (.002)	-.006 (.004)	.065 (.018) <sup>***</sup>	-334.179	684.358	712.561
Preferred model (AIC and BIC)	.092 (.128)	.126 (.051) <sup>*</sup>	-.000 (.004)	.017 (.002) <sup>***</sup>	-.001 (.001)	←	-.006 (.004)	.065 (.018) <sup>***</sup>	-334.513	683.025	707.703
<b>Genetic stability</b>											
Mean effect size	.844 (.038) <sup>***</sup>	—	—	—	—	—	—	.119 (.021) <sup>***</sup>	-340.007	684.014	691.065
Full model	.518 (.123) <sup>***</sup>	.030 (.043)	.014 (.005) <sup>**</sup>	.017 (.003) <sup>***</sup>	-.002 (.001) <sup>*</sup>	-.001 (.001)	-.013 (.006) <sup>*</sup>	.036 (.010) <sup>***</sup>	-287.580	591.159	619.363
Preferred model (AIC and BIC)	.553 (.024) <sup>***</sup>	→	→	.016 (.001) <sup>***</sup>	-.001 (.001)	←	-.013 (.006) <sup>*</sup>	.034 (.009) <sup>***</sup>	-287.779	585.558	603.185
<b>Environmental stability</b>											
Mean effect size	.397 (.041) <sup>***</sup>	—	—	—	—	—	—	.136 (.023) <sup>***</sup>	-364.406	732.813	739.864
Full model	-.275 (.146)	.189 (.060) <sup>**</sup>	.003 (.006)	.010 (.004) <sup>**</sup>	.004 (.002) <sup>**</sup>	-.007 (.002) <sup>***</sup>	-.016 (.007) <sup>*</sup>	.086 (.020) <sup>***</sup>	-315.639	647.279	675.482
Preferred model (AIC and BIC)	-.253 (.143)	.176 (.055) <sup>**</sup>	.006 (.001) <sup>***</sup>	←	←	-.007 (.001) <sup>***</sup>	-.015 (.007) <sup>*</sup>	.086 (.020) <sup>***</sup>	-316.147	644.295	665.447
<b>Corrected environmental stability</b>											
Mean effect size	.652 (.064) <sup>***</sup>	—	—	—	—	—	—	.215 (.044) <sup>***</sup>	-379.335	762.669	769.712
Full model	-.656 (.270) <sup>*</sup>	.440 (.045) <sup>***</sup>	-.017 (.010)	.023 (.007) <sup>***</sup>	.003 (.002)	-.001 (.001)	-.024 (.010) <sup>*</sup>	.102 (.045) <sup>*</sup>	-328.563	673.126	701.298
Preferred model (AIC and BIC)	-.654 (.270) <sup>*</sup>	.440 (.107) <sup>***</sup>	-.018 (.010)	.024 (.006) <sup>***</sup>	.002 (.001)	←	-.023 (.010) <sup>*</sup>	.102 (.045) <sup>*</sup>	-328.748	671.495	696.145

*Note.* Dashes indicate that the parameter was not estimated.  $b_0$  represents the intercept,  $b_{0-3}$  represents the coefficient for the first linear segment,  $b_{3-15}$  represents the coefficient for the following segment, and so on until  $b_{60-90}$ . For simplified models, arrows are used to indicate the parameter estimate that applies during the given developmental period.  $\tau$  represents the random-effect standard deviation in the between-study effect size. AIC = Akaike information criterion; BIC = Bayesian information criterion; LL = loglikelihood.

<sup>\*</sup>  $p < .05$ . <sup>\*\*</sup>  $p < .01$ . <sup>\*\*\*</sup>  $p < .001$ .

environmental stability, on the other hand, continues to rise throughout the life span. At age 15, the true environmental correlation across time is expected to be .52, and at age 30, it is expected to be .69. True environmental stability continues to rise at age 45 (expected  $rE = .81$ ) and age 60 (expected  $rE = .88$ ) and does not reach the predicted asymptote by age 90 (expected  $rE = .96$ ).

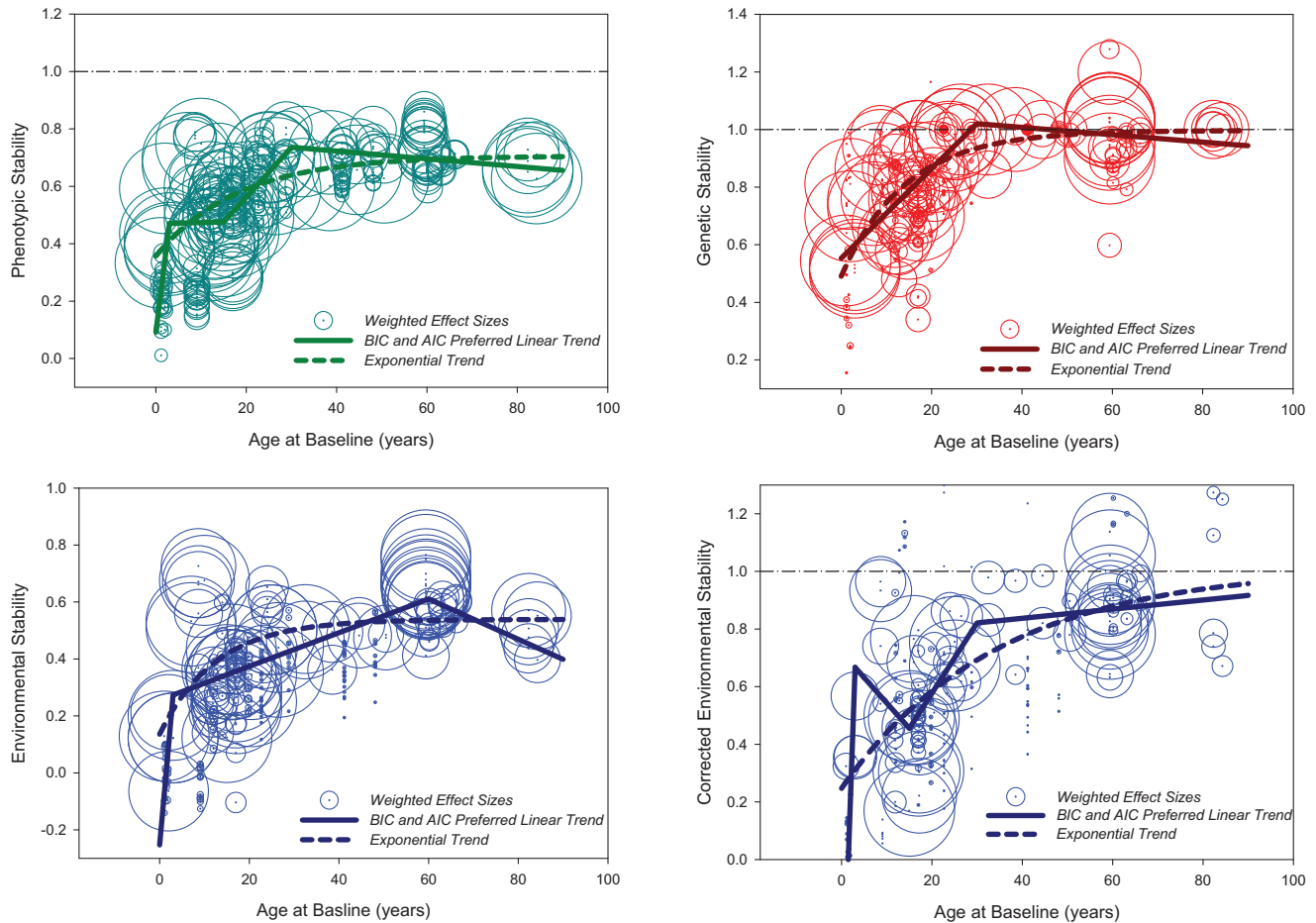
These trends are readily apparent from Figure 4. The empirical story is the same across linear, nonlinear, and graphical representations of the data. Phenotypic stability is very limited in infancy but increases fairly quickly over early development and adolescence. Increases in phenotypic stability plateau near age 30 and remain at this level. Genetic stability is substantial even in infancy, but this type of stability also increases steadily over childhood and adolescence to reach a plateau at unity near age 30. Environmental stability rises sharply in infancy and then slowly throughout adolescence and adulthood. Corrected environmental stability rises at a steady rate throughout infancy, childhood, and adulthood, continuing even into old age. Slight differences between the connected linear and exponential trends are apparent for phenotypic stability (e.g., the pause in increasing stability during childhood found in the linear model) and corrected environmental stability (e.g., the jagged increase and decrease of stability found in childhood).

**Age Trends in the Contribution of Genes and the Environment to Stability**

Table 6 presents the age trends in the genetic and environmental contribution to phenotypic stability. For the linear models, the BIC comparisons indicate that including age as a moderator of the genetic contribution to stability actually reduces model parsimony without a compensatory increase in model fit according. In other words, genetic effects exert a constant, moderate effect ( $b_0 = .358$ ,  $p < .001$ ) on phenotypic stability across the life span. On the other hand, environmental contributions to stability vary with age and are best approximated by a model with slopes before and after age 30. In very early childhood, the environment does not contribute to phenotypic stability ( $b_0 = .034$ ,  $p = .09$ ), but the environmental contribution increases until age 30 ( $b_{0-30} = .008$ ,  $p < .001$ ) and plateaus afterward ( $b_{30-90} = .001$ ,  $p = .21$ ). The AIC preferred models contain substantially more complexity in terms of the number of free slopes that are required. However, visual inspection of Figure 5 indicates that the BIC and AIC preferred trend lines largely overlap. For the genetic contribution, the AIC preferred model implies that the genetic contribution vacillates around the constant value implied by the BIC model. In infancy and adolescence, the estimate of the genetic contribution is slightly lower, and in childhood and adulthood, the estimate is slightly higher. For the environmental contribution, the trend lines overlap nearly perfectly, except in old age, where the AIC preferred model indicates a slight decline.

Table 4 presents the continuous exponential results. The exponential models closely align with the connected-linear models. The continuous function indicates an upper asymptote of .38 for genetic contributions to stability and .33 for environmental contributions to stability. For the genetic contribution, this asymptote is reached very early in life, by age 8. In comparison, the increase in the environmental contribution to stability occurs slowly throughout the life span. At age 15, the expected environmental contribution is only .17 correlation units, and it continues to .25 by age 30





**Figure 4.** Age trends in phenotypic, genetic, environmental, and measurement error corrected environmental stability assuming a 5.56-year time lag between assessments. Circles surrounding data points are scaled by the weighting variable (described in the Analytic Approach section in the text) such that larger circles carried *more* weight in the analysis. Some data points were estimated to be out of bounds of the logical limit of a correlation (i.e.,  $-1$  to  $1$ ). This likely results from parameter imprecision and slight violations of the traditional assumptions of behavioral genetic models (e.g., monozygotic twins correlated more than twice as strongly as dizygotic twins). A total of 13 such estimates are not displayed on the graph for corrected environmental stability because they were outside of the viewable range. AIC = Akaike information criterion; BIC = Bayesian information criterion.

and continues to rise past age 60 (expected environmental contribution = .31).

The identified linear and nonlinear trends are apparent in Figure 5, with reference to the expected exponential life-span trend for phenotypic stability plotted in green. In infancy, phenotypic stability equals the genetic contribution to stability, and genetic effects exert an impressive and almost constant influence on stability across the entire life span. However, increasing environmental stability, from negligible in childhood to almost equivalent importance in old age, is entirely responsible for *increasing* phenotypic stability. This trend is consistent with that found in Figures 3 and 4. Genetic influences increase in stability across development, but the total variation in personality associated with genetic differences decreases across the life span. This tradeoff results in a nearly constant genetic contribution to phenotypic stability. Both environmentality and environmental stability increase across the life span. Thus, the combination of environmental effects persist-

ing to later ages to a greater extent and accounting for personality variance to a greater extent results in an increasing environmental contribution to phenotypic stability across the life span.

### Moderation Analysis

The goal of the next phase of the analysis was to determine if effect sizes differ as a function of moderators other than age and time lag. We accomplished this goal by adding the effects coded moderators as predictors in the exponential model for each outcome. We chose to use the exponential model for each outcome to present a standard set of results and because the exponential model tended to overlap substantially with connected-linear trends. The use of effects coded variables allows the coefficients to be interpreted as deviations from the average trend (i.e., that reported in Table 4) rather than deviations from a reference category (e.g., the trend for extraversion).

Table 6  
Age Trends in the Genetic and Environmental Contribution to Phenotypic Stability

Regression model	$b_0$	$b_{0-3}$	$b_{3-15}$	$b_{15-30}$	$b_{30-60}$	$b_{60-90}$	$b_{\text{time lag}}$	$\tau$	LL	AIC	BIC
<b>Genetic contribution to stability</b>											
Mean effect size (BIC preferred model)	.358 (.017) <sup>***</sup>	—	—	—	—	—	—	.053 (.016) <sup>**</sup>	-.282.281	568.561	575.604
Full model	.139 (.086)	.090 (.029) <sup>**</sup>	-.008 (.002) <sup>***</sup>	.008 (.002) <sup>***</sup>	-.002 (.001)	-.001 (.001)	-.001 (.005)	.050 (.016) <sup>**</sup>	-.275.728	567.456	595.627
Preferred model (AIC)	.139 (.086)	.090 (.029) <sup>**</sup>	-.008 (.002) <sup>***</sup>	.007 (.002) <sup>***</sup>	-.002 (.001) <sup>*</sup>	←	-.001 (.005)	.050 (.016) <sup>**</sup>	-.275.778	565.556	590.206
<b>Environmental contribution to stability</b>											
Mean effect size	.175 (.024) <sup>***</sup>	—	—	—	—	—	—	.062 (.011) <sup>***</sup>	-.330.870	665.740	672.783
Full model	-.082 (.064)	.053 (.022) <sup>*</sup>	.006 (.002) <sup>**</sup>	.008 (.002) <sup>***</sup>	.002 (.001) <sup>***</sup>	-.003 (.001) <sup>*</sup>	-.008 (.003) <sup>*</sup>	.019 (.007) <sup>*</sup>	-.275.956	567.911	596.083
Preferred model (AIC)	-.079 (.065)	.051 (.023) <sup>*</sup>	.007 (.001) <sup>***</sup>	←	.003 (.001) <sup>***</sup>	-.003 (.001) <sup>***</sup>	-.008 (.004) <sup>*</sup>	.019 (.007) <sup>*</sup>	-.276.004	566.008	590.659
Preferred model (BIC)	.034 (.020)	→	→	.008 (.001) <sup>***</sup>	.001 (.001)	←	-.006 (.003)	.024 (.007) <sup>**</sup>	-.279.846	569.692	587.299

Note. Dashes indicate that the parameter was not estimated.  $b_0$  represents the intercept,  $b_{0-3}$  represents the coefficient for the first linear segment,  $b_{3-15}$  represents the following segment, and so on until  $b_{60-90}$ . For simplified models, arrows are used to indicate the parameter estimate that applies during the given developmental period.  $\tau$  represents the random-effect standard deviation in the between-study effect size. AIC = Akaike information criterion; BIC = Bayesian information criterion; LL = loglikelihood. \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

**Moderation by Big Five traits.** Table 7 presents the moderation results for the Big Five traits. In general, very few effect sizes differed by trait category. Where there were statistically significant differences, the magnitude of the differences tended to be very small. For example, extraversion and conscientiousness are significantly more heritable than the average personality outcome, and agreeableness is significantly less heritable. However, these deviations amount to about a difference of 1% of variance. Conscientiousness displayed the most consistent deviations from the average trend, as this trait tends to be more environmental and more stable phenotypically and environmentally (but not genetically), with both genes and the environment contributing more to its stability (as would be expected since it is overall more phenotypically stable). Extraversion tends to be influenced more by genes and is more stable phenotypically and genetically, with genes contributing more to stability. Agreeableness tends to be more environmental and less stable phenotypically and environmentally, and genes contribute less to stability. No significant differences were found for neuroticism or openness. Despite several statistically significant differences between the Big Five traits, the major conclusion is that differences are strikingly trivial.

**Moderation by self-versus other report.** Table 8 presents the moderation results for self-compared to other report assessments of personality, and again, very few differences were found. Assessments using self-report tended to be less heritable and more environmental. The stability of environmental effects was lower, as were both genetic and environmental contributions to stability. Each effect size is likely trivial for pragmatic purposes.

**Moderation by broad versus narrow measurement.** Table 8 additionally presents the moderation results for broad compared to narrow measurement of personality. Broad measures tended to be less heritable and more environmental. They tended to be less stable phenotypically but more stable in terms of corrected environmental stability. Environments contributed to stability to a lesser extent for broad measures. Similar to the previous moderation analyses, the differences are of a trivial magnitude.

**Sensitivity Analysis: Report Format**

Differences in report format are heavily clustered at certain ages (i.e., informant report predominates at early ages and self-report in adulthood). Although our above moderation analysis indicated that effect sizes did not substantially differ by report format, we were interested in whether the age trends identified earlier would hold when only self-report effect sizes were used. This approach avoids potential differences in effect sizes on the basis of report format being mistaken for age trends in effect sizes, at the cost of reducing the meta-analytic sample size and removing information derived from infant and early childhood samples. Because the exponential and spline models applied in our main analyses largely agreed with one another, we focused this sensitivity analysis on the exponential model, which we fit only to effect sizes derived from self-report data. Results are tabulated in Table S1 and illustrated in Figures S1–S3 in the online supplemental materials. Note that age 9 years is the youngest self-report effect size, and we do not extrapolate the expected trend line to younger ages.

Consistent with the results of analyses of the complete data set, results of analyses of only self-report data indicate that

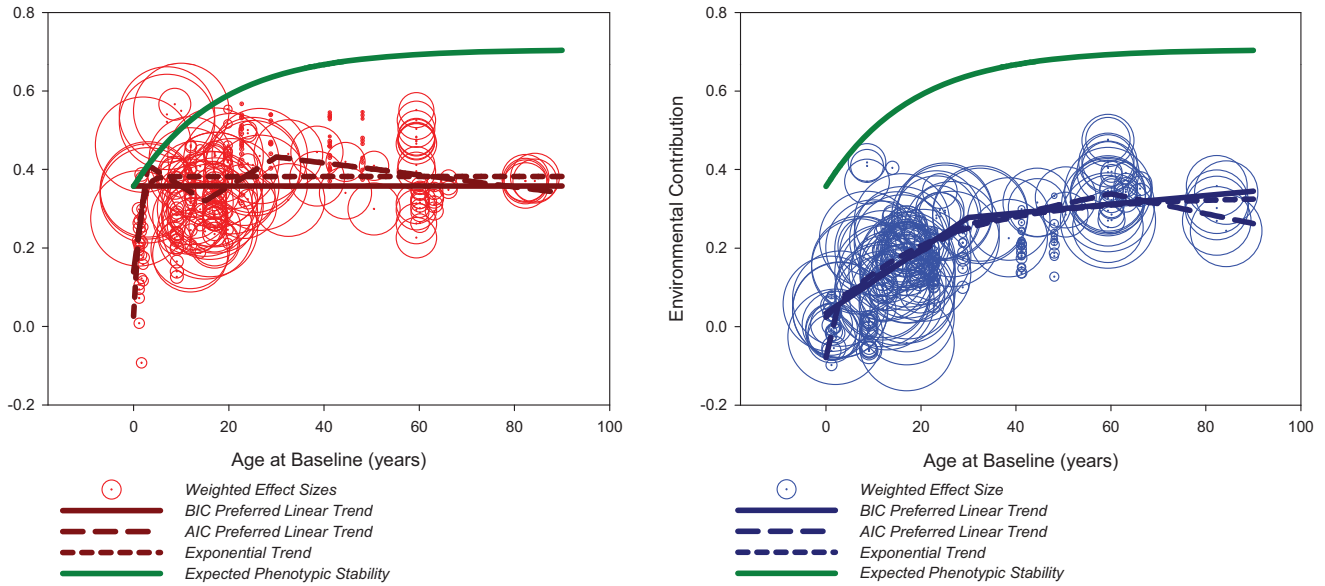


Figure 5. Age trends in genetic and environmental contributions to phenotypic stability assuming a 5.56-year time lag between assessments. Circles surrounding data points are scaled by the weighting variable (described in the Analytic Approach section in the text) such that larger circles carried more weight in the analysis. AIC = Akaike information criterion; BIC = Bayesian information criterion.

phenotypic, genetic, and environmental stabilities increase significantly with age, particularly up until about age 30 years (see Figure S2 in the online supplemental materials). In fact, these rises in stability are somewhat more pronounced for the analyses of exclusively self-report data than for the analyses of the complete data set, indicating that they are not artifacts of blending information from self- and informant report formats. Age-related trends for heritability and environmental stability derived from analyses of only self-report data (see Figure S1 in the online supplemental materials) are weaker than for those derived from analyses of the complete data set. For analyses of only self-report data, heritability does appear to decline (and environmental stability to increase) slightly early in life, but this trend is not statistically significant. The predominant pattern is for constant magnitudes of genetic and environmental influences

on personality across the life span. Corrected environmental stability did significantly increase with age, but this increase was modest and plateaued quickly. Of course, if the most dramatic changes in heritability and environmental stability primarily occur in the first decade of life, then these analyses of only self-report data (which do not contain data from infancy and early childhood) may not be well equipped to detect them. Finally, consistent with results of analyses of the entire data set, analyses of only self-report data indicate that age-related increases in phenotypic stability are predominantly attributable to increases in environmental contributions (see Figure S3 in the online supplemental materials). Analyses of only self-report data do indicate slight increases in genetic contributions with age. However, environmental contributions are still the predominant contributor to increasing phenotypic stability. Starting at age 9, the genetic

Table 7  
Moderation of Effect Sizes by the Big Five

Outcome	E	A	C	N	O
Heritability	.011 (.004)*	-.009 (.003)**	.005 (.001)***	.001 (.003)	-.008 (.005)
Environmental stability	-.010 (.004)*	.016 (.009)	-.004 (.001)***	.000 (.003)	-.002 (.009)
Corrected environmental stability	-.001 (.003)	.011 (.007)	-.004 (.001)***	.002 (.004)	-.008 (.006)
Phenotypic stability	.003 (.001)**	-.009 (.004)*	.007 (.002)***	.000 (.002)	-.001 (.005)
Genetic stability	.004 (.002)*	-.008 (.005)	.001 (.001)	-.002 (.002)	.005 (.003)
Environmental stability	.000 (.002)	-.003 (.005)	.011 (.002)***	-.004 (.004)	-.004 (.012)
Corrected environmental stability	.004 (.004)	-.007 (.002)**	.021 (.009)*	-.011 (.008)	-.007 (.007)
Genetic contribution to stability	.007 (.002)**	-.010 (.003)***	.006 (.001)***	.000 (.002)	-.003 (.005)
Environmental contribution to stability	-.001 (.001)	-.002 (.003)	.001 (.000)*	-.001 (.001)	.002 (.003)

Note. All models control for exponential age trends and linear time-lag trends in effect sizes. The Big Five variables were effects coded (see the Method section in the text), and therefore, the parameter estimates represent deviations from the average trend. E = extraversion; A = agreeableness; C = conscientiousness; N = neuroticism; O = openness.  
\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .



Table 8  
Moderation of Effect Sizes by Report Format and  
Trait Bandwidth

Outcome	Self- (.5) vs. informant (-.5) report	Broad (.5) vs. narrow (-.5) measurement
Heritability	-.007 (.001)***	-.008 (.002)***
Environmental stability	.007 (.001)***	.007 (.002)***
Corrected environmental stability	.008 (.001)***	.008 (.001)***
Phenotypic stability	-.006 (.004)	-.005 (.002)*
Genetic stability	-.001 (.002)	-.001 (.002)
Environmental stability	-.020 (.002)***	-.014 (.007)
Corrected environmental stability	-.033 (.015)*	.018 (.008)*
Genetic contribution to stability	-.011 (.002)***	-.005 (.004)
Environmental contribution to stability	-.003 (.000)***	-.003 (.001)**

Note. All models control for exponential age trends and linear time-lag trends in effect sizes.

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

contribution rises by .14 correlation units until it plateaus, but the environmental contribution rises by .31 correlation units.

### Sensitivity Analysis: Publication Bias

We evaluated the robustness of our results to possible publication bias by using funnel plots to guide a series of sensitivity analyses. Funnel plots involve plotting the outcome effect sizes against its imprecision (standard error). Estimates with larger standard errors (e.g., those derived from small-sample studies) are expected to have greater spread from the meta-analytic expectation, whereas estimates with small standard errors (e.g., those derived from large-sample studies) are expected to have relatively small deviations from the meta-analytic expectation. Thus, a funnel shape is expected for this type of plot, with the amount of spread of scatter continuously related to the larger standard error studies. Importantly, regardless of spread, the scatter should be symmetrical around the meta-analytic expectation. Where effect sizes are conspicuously absent from the funnel such that the spread is asymmetrical, publication bias is possible.

To produce our funnel plots, we residualized the effect sizes based on the expected age and time-lag effects from the continuous exponential models. In instances in which the standard errors were positively skewed (which was the case for phenotypic stability, genetic stability, and corrected environmental stability), we log transformed the standard errors to aid in visualization. Two trends were particularly apparent.

First, there was evidence of effect sizes that fell substantially outside of the expected funnel distribution for heritability, environmental stability, and corrected environmental stability. For heritability, the distribution of residual effect sizes was negatively skewed with a few highly negative values (for environmental stability, the distribution of effect sizes was positively skewed). These effect sizes were primarily drawn from the first years of life. Returning to Figure 3, it is apparent that there are several effect sizes that indicate low heritability during this period of the life span. However, all of these studies are very low precision, and the estimated trend line reflects several high-precision, high-heritability estimates. There are a few possible explanations. If heritability approaches the

upper limit of 1.0 near birth, then it is not logically possible for deviations to occur such that heritability accounts for more than 100% of the variance. Rather, the small-sample and low-precision studies will be more likely to estimate heritability to be lower than the high-precision studies due to the greater amount of sampling variability. Alternatively, it may be the case that the low-precision estimates are indicative of substantial heterogeneity of effect sizes. Our moderation analysis was unlikely to detect this heterogeneity on the basis of the measured moderators as the data points carried little weight in the analysis. It may also be the case that the heterogeneity is due to moderators that we did not choose to evaluate. More high-precision studies of early child development would be necessary to evaluate this issue with greater accuracy.

Second, potential evidence of publication bias was found for heritability, environmental stability, corrected environmental stability, and genetic stability in that a symmetrical funnel shape was not found for low-precision studies. To evaluate the influence that such possible publication bias had on the results, we deleted any effect size that had standard errors greater than .075, which corresponded to where the asymmetry became apparent. Then, we computed the parameters for the exponential model based on this limited data set. The life-span trends from the resulting model were essentially the same as those reported in Table 4. Therefore, although there may be some publication bias for low-precision studies, the meta-analytic estimates reported in this article can be considered robust to this bias.

### Discussion

The current project is the first comprehensive quantitative review of genetic and environmental mechanisms of differential stability of personality across the life span. Replicating previous findings (Ferguson, 2010; Roberts & DelVecchio, 2000), trait stability increased monotonically until age 30 with a possible stall in early childhood. We, however, found no clear evidence for continued increases in phenotypic stability, a pattern that is consistent with some previous research (Ferguson, 2010; Terracciano et al., 2006) and inconsistent with other studies (Lucas & Donnellan, 2011; Roberts & DelVecchio, 2000). Genetic stability increased from moderate in infancy to near perfect by age 30 and remained near perfect across adulthood. Environmental stability (uncorrected), in contrast, displayed almost complete instability in childhood but increased to about half as stable as genetic influences by adolescence. Correcting for measurement error, environmental stability was weak in early childhood, increased with age, and peaked at a level only slightly less than that of genetic stability. Additionally, we found that genetic influences contributed to phenotypic stability at a relatively stable rate. In contrast, environmental contributions to stability changed substantially with age and accounted for the majority of increasing phenotypic stability in both the full and self-report analyses. By midlife, genetic and environmental effects were found to contribute almost equally to phenotypic stability. Effect sizes were very similar across different personality traits and measurement paradigms.

Our meta-analysis is particularly innovative in its ability to draw inferences about the developmental trends in the genetic and environmental sources of stability across close to the entire life span. Previous studies have provided insights into specific periods of development, but no individual study has been able to plot

trends in genetic and environmental stability from birth to near the end of the life span. By combining studies of circumscribed periods of development, we were able to make several novel insights into personality development across much longer periods than those examined in the individual studies. Phenotypic, genetic, and environmental stabilities of personality increase substantially in the first 3 decades of life. Likely the most innovative and surprising finding of the current analysis is that the genetic contribution to stability remains relatively constant across the life span compared to the large increases in environmental contributions to phenotypic stability.

### Theoretical Implications

We detailed hypotheses from intrinsic and social maturation perspectives, some of which our results confirmed and some of which they did not. The FFT correctly predicted that phenotypic and genetic stability would peak near age 30 and that genes would represent the primary contribution to stability across the life span. Discrepancies with the FFT were found in the increasing stability of and variance accounted for by environmental effects. As originally reported in a classic cross-sectional meta-analysis of twin studies (McCartney, Harris, & Bernieri, 1990), heritability of personality decreased with age, and environmental influences increased. Moreover, environmental influences increased in stability across development, evincing impressive levels in adulthood. Furthermore, the increase in phenotypic stability could largely be explained by increasing environmental contributions. These results of increasing environmentality, increasing environmental stability, and increasing contributions of environmental factors to phenotypic stability were predicted by the SPMs. Deviations from SPMs were observed in the near perfect stability of genetic effects after age 30. This suggests that the environmental changes typically experienced in adulthood do not meaningfully influence gene expression relevant to personality.

Mechanisms of Gene  $\times$  Environment interaction and gene-environment correlation may be consistent with our results. In early childhood, it may be the case that substantial and extensive Gene  $\times$  Environment interaction takes place based on the shared environment provided by parents. This process would explain the very large heritability of personality in early childhood. As children mature and begin to leave the shared environment, heritability begins to decrease indicating that the Gene  $\times$  Environment interactions may not be particularly lasting (i.e., a scarring effect that would persist beyond the environmental exposure) but rather may frequently recur over early development. However, one would expect that this process would entail a decrease in genetic stability, as children move away from shared environments to which they differentially respond. We found the opposite trend, a fairly constant increase in genetic stability through childhood. Gene-environment correlation may explain the increase in genetic stability. Children increasingly gain the autonomy to select environments consistent with their genotype with age (Scarr & McCartney, 1983). Age-dependent gene expression is another possible explanation for relatively low genetic stability in childhood with increases into adulthood. As children mature, genetic effects may activate, while others deactivate over child development but remain in relatively consistent activation states across adulthood. Because genetic effects decrease in importance but also increase in

stability, it is possible that any or all of these mechanisms occur simultaneously.

Gene  $\times$  Nonshared Environment interaction may contribute to the life-span trends for environmental influences on personality. If children encounter more novel experiences with age and differentially react to the experiences on the basis of genotype, then this process could explain the increasing influence of environmentality. If these experiences relate to one's occupational, interpersonal, or romantic identity, then it is likely that these relatively enduring environments would cause an increase in environmental stability as well.

Explanations relying on complex gene-environment dynamics should be weighed against more parsimonious theories that postulate primarily direct genetic and environmental effects. For example, SPMs argue that people accumulate environmental experiences relevant for their personal identity with age and that these experiences have a causal impact on personality development. As these unique experiences add up over development, genetically identical individuals come to resemble each other less and less. In the context of a behavioral genetic model, this process means that heritability would decrease with age and that environmentality would increase. Genetic effects may not be perfectly stable from birth because the accumulating environmental experiences act as a social push and thus constrain phenotypic expression to a certain extent. As individuals increasingly create a stable identity or environmental niche, they encounter less novel environments (i.e., increasing environmental stability), and this mechanism also acts to reduce novel social pushes (i.e., increasing genetic stability).

What evidence is there to prefer the more complex explanation involving gene-environment dynamics over independent and additive effects of genes and the environment? As discussed in the introduction to this article, empirical examples of gene-environment dynamics are too common to ignore (see Dick, 2011; Johnson, 2007; Tabery, 2007). Gene  $\times$  Environment interactions "are ubiquitous in nature" and evolutionarily adaptive (Penke, Denissen, & Miller, 2007, p. 578). Phenotypic plasticity in response to environmental circumstances and resources allows organisms to adjust to environmental variation (Dall, Houston, & McNamara, 2004). This feature may explain the extremely high heritability of personality in very early childhood. During this period, the ultimate phenotype is very amenable to environmental inputs, but the phenotype responds differentially on the basis of genotype. As this extreme plasticity subsides, other etiological mechanisms may begin to emerge as impactful on personality development. Future research that documents gene-environment interplay as well as direct genetic and environmental effects will be important to more precisely assess the magnitude of the importance of the various processes.

### Current and Future Directions in Behavioral Genetic Work

Despite the theoretical progress reviewed above, it has been notoriously difficult for behavioral genetic analyses to uncover environments that have an impact on psychological outcomes (Plomin & Daniels, 1987). For example, Turkheimer and Waldron (2000) found that less than 5% of the variance attributed to the nonshared environment could be accounted for by measurable aspects of the environment. This finding led Turkheimer (2000) to

affirm the gloomy prospect that nonshared environmental effects were too idiosyncratic, complex, or transient to identify with scientific inquiry. McCrae and Costa (2008) similarly commented that “ambitious attempts to pin down substantive contributions of the nonshared environment have largely failed” (p. 168). Importantly, our findings indicate that a substantial proportion of variance in personality traits is influenced by environmental experiences acting to make siblings different from one another that is not attributable to measurement error and is stable across time. Although we take seriously previous failures to pin down measurable nonshared environmental experiences, our findings suggest a substantial amount of variance exists in personality traits that is truly environmentally mediated, not random error of measurement, and unrelated to state-like fluctuations. This result should be considered an encouraging prospect for future investigation of measurable environments. Longitudinal, genetically informative samples will likely be necessary to pinpoint the specific enduring environments and locations in development during which the environment exerts a causal effect. Whether such measurable environments can be realistically expected to explain the totality of stable nonshared environmental effects is, however, unclear.

The *missing heritability* problem is a similar unresolved issue in the behavioral genetic literature (Maher, 2008). Just as researchers have largely been unable to find measured environments that can account for substantial amounts of the latent environmental influences found in twin models, measured genes have been found to account for only a minute portion of variance in personality, with very few replicable genetic markers (Terracciano, Sanna, et al., 2010). Establishing an array of genetic variants that reliably predict personality variation would be strong support for biological models of personality. Again, the current results indicate that lasting genetic influences affect personality. The prevailing perspective is that thousands of genes have an infinitesimal influence on complex phenotypes, but with large enough sample sizes, these effects should be able to be reliably detected (Plomin, 2013). However, Vinkhuyzen et al. (2012) used genome-wide complex trait analysis and were able to account for only about a third of the missing heritability for personality. This method is considered assumption free in that unrelated individuals are compared to one another on the basis of similarity among measured genes, and therefore, it provides a corroborating estimate of additive heritability. According to this study, two thirds of the variance traditionally assumed to be of genetic origin is due to rare genetic variants not captured by current genotyping technology, nonadditive genetic effects (i.e., dominant genes, epistasis), dependencies between genes and environments, or other factors that would raise estimates of heritability in twin and family studies. Therefore, the totality of genetic influences on personality is unlikely to be found using common molecular genetic approaches. Nonadditive genetic effects on personality are real possibilities that should be considered in ongoing work (see Keller, Coventry, Heath, & Martin, 2005; Rettew, Rebollo-Mesa, Hudziak, Willemsen, & Boomsma, 2008; Verweij et al., 2010). We were, however, unable to evaluate such nonadditive effects, as many of the primary studies included in our meta-analysis only reported results from models of additive genetic variation. As more evidence of nonadditive genetic effects accumulates, the trends in heritability, genetic stability, and genetic contributions to stability may need to be updated to evaluate

whether narrow, additive genetic effects and dominant genetic effects have different life-span trends.

The lack of identified measured environments or genes that are influential for personality development may be potentially related. If gene–environment interplay explains the developmental trajectory of the differential stability of personality, then it is unlikely that specific candidate genes or candidate environments would be directly linked to phenotypic variation. Again, the failures to pin down environmental effects and the problem of missing heritability seem to imply that dynamic processes undergird personality development rather than direct genetic or environmental influences. We have highlighted some of the promising new research in this area, but there remain many unknowns with regard to what environments are influential, when, and for whom. It seems that it will be important to investigate chosen environments (e.g., niche building), dyadic relationships (e.g., peer and parent relationships), and discrete experiences (e.g., stressful life events). Similarly, it will be important to identify when in development and in what environments genetic influences on personality are activated and expressed. Early childhood and adolescence appear to be a time of relative genetic instability, with genetic effects decaying in magnitude. Do direct, social pushes or constraints on patterns of behavior limit the ability of individuals to act according to their genetic predispositions as they develop? Or do individuals increasingly respond to unique life experiences differently on the basis of their genotype? To resolve these questions, research will need to take into account not only environmental experience but the unique reaction that each individual will have to different sorts of environments.

We view longitudinal, genetically informative studies that assess individuals’ personal characteristics and preferences, plausible biological endophenotypes of personality, the social environment, the macroenvironmental resources available, and important life events as crucial to the success of ongoing personality research. Although a number of empirical examples of gene–environment interplay can be found in the literature, longitudinal extensions of these designs are infrequent. Are Gene  $\times$  Environment interactions lasting or fleeting? Are they only active during certain critical periods of development? Addressing these questions requires not only identifying gene–environment interplay but also tracking the effects through time (e.g., Rhemtulla & Tucker-Drob, 2012).

## Strengths and Limitations

This study is the first to empirically demonstrate the genetic and environmental stability trends of personality across nearly the entire life span. By aggregating data and applying meta-analytic techniques, we have provided a provocative picture of personality development. We obtained data from 24 studies that sampled more than 40,000 individuals who ranged in age from infancy to old age. Our approach, however, was not without limitations. First, we chose to apply random-effects connected-linear spline models to describe the age trends in the data. These models have the advantage of being able to detect discontinuities in trends. For example, we found that increasing phenotypic stability appears to stall in early childhood. There are two interpretations of this finding. Either phenotypic stability truly stalls during this developmental period or the model was overspecified and simply picked up on



noise in the data that was not meaningful. To counteract this possibility, we have provided results from a number of different modeling approaches. By providing the most complex connected-linear spline model, the trends can be examined with the greatest flexibility. Alternatively, the continuous exponential trend provides the best general impression of the data that is potentially less influenced by noise. Importantly, visual inspection of the trend lines indicates that every model tells essentially the same story with only slight deviations.

Second, this meta-analysis is somewhat unique in the sense that rather than obtaining reported effect sizes for an outcome of interest, we obtained summary data of different group correlation matrices. This level of data extraction allowed us to fit a standard model to the data derived from all studies such that effect sizes could be meaningfully aggregated. However, this approach has some drawbacks. For example, reporting of full cross-sibling cross-time correlation matrices was rather rare. Instead, we were required to derive these matrices from a variety of different behavioral genetic models reported in the original articles. When fully saturated models were reported, the parameters allowed for a direct estimation of the underlying data. When only reduced or trimmed models were reported, the group correlations were imperfectly estimated. In practice, the majority of studies reported models that had trimmed the shared environmental parameters. Short of having access to all of the raw data from the studies used, we have provided the most accurate estimates possible from published data. A further issue associated with extracting correlation matrices relates to differences in variance across time points. Access to covariance, rather than correlation, matrices would have allowed for an estimate of whether total, genetic, or environmental variance components change with age.

Third, we were able to examine stability for individual developmental periods and whether differences were found for several moderators. Every effort was made to code effect sizes in the most straightforward way based on empirical associations that have been reported between different assessment strategies. However, some choices were made based on our best judgment. For example, the chosen developmental breaks at the ages of 3, 15, 30, and 60 are somewhat arbitrary. Shifting these values may have produced different results as this would reorganize the distribution of effect sizes. We would point to the substantial overlap of the preferred connected-linear models with that found with the exponential model as evidence that our break points correctly split the data.

Fourth, the moderation analysis may have been underpowered if there was not sufficient data density across levels of the moderator for the entire life span. This limitation may be particularly relevant for self- compared to informant report format as parent report is primarily used in early childhood and self-report is primarily used in adulthood. Although we found very consistent results with respect to age trends in stabilities and contributions to stability when analyses were restricted to effect sizes from only self-report data, the age trends in heritability and environmentality appeared weaker when only self-report data were used. If the most dramatic changes in heritability and environmentality primarily occur in the first decade of life, then analyses of only self-report data (which do not contain data from infancy and early childhood) may not be well-equipped to detect them. However, it is also possible that parents tend to exaggerate differences between their (nonmonozy-

gotic) children, which acts to magnify heritability estimates in early childhood (Saudino & Cherny, 2001). For example, Saudino (2012) found that the heritability of activity level at age 2 was .84 for parent report, but observer report was only .55. Unfortunately, Saudino's is the only study of early childhood included in the meta-analysis that used both parent and observer report. Importantly, there are multiple sources of accurate personality-relevant information (Funder, 1999). These sources of information are differently available to the self (e.g., internal thoughts and feelings), close relatives (e.g., observation of a wide range of behavior from birth), peers (e.g., observation of behavior in an important social context), and trained observers (e.g., observation of a narrow range of behaviors in a controlled environment). Using multiple assessment strategies provides a more complete picture of personality, but there are obvious limitations. Self-report measures are not possible to collect in infants and very young children, and even in middle childhood and early adolescence, the psychometric properties (e.g., reliability, acquiescence, factor coherence, and factor differentiation) and validity of self-report data are less established and at times questionable (Soto, John, Gosling, & Potter, 2008; Tackett et al., 2012). Bridging research across infant, child, adolescent, and adult development using metrically invariant measurement approaches that capitalize on the unique advantages of multiple sources of personality-relevant information should be a goal for future research.

Fifth, we used Cronbach's alpha as an estimate of internal consistency in order to correct the estimates of environmentality and environmental stability for measurement error. However, Cronbach's alpha is an overestimate of measurement error in several situations, particularly when the measure is multidimensional (Schmitt, 1996). Thus, our models may have overcorrected for measurement error. With this limitation in mind, the corrected estimates should be interpreted as lower bounds of environmentality and upper bounds of environmental stability.

Sixth, we focused on the effect of participant age on the outcomes of interest. We controlled for linear differences in time lag, but it is also possible that stability decays in a nonlinear manner as more time passes between assessments toward some lower asymptote. As a sensitivity analysis (not presented), we included a quadratic time-lag term in each of the continuous exponential models. This addition did not substantively change any of the life-span trends, and the quadratic term tended to be small in magnitude.<sup>8</sup> Although the current article is specifically focused on age-based life-span trends in stability, it will be valuable for future research to examine time-based trends in stability. This analysis will help to determine the most appropriate functional form of stability decay and whether such decay interacts with the age of the participants (e.g., stability may fall precipitously at early ages but persist through time in adulthood).

Finally, the effect sizes were derived from published studies that assessed individuals from different cohorts that grew up in different periods. It may be the case that the genetic and environmental influences on personality not only depend on the chronological age of an individual but also depend on the macroenvironmental con-

<sup>8</sup> In fact, the average absolute deviation of the age-related parameters from a model that included a quadratic time effect and those reported in Table 4 was only .016.



text that changes across cohorts or periods (e.g., Heath et al., 1985). Identifying these types of temporal trends would be highly novel research for personality psychology. In the context of the current study, we assume that age differences across longitudinal studies converge to provide a reasonable viewpoint of the larger life-span trend (Bell, 1953). The alternative to this limiting assumption is to track a single cohort through the life span to examine the differential stability of personality, but no such sample is available.

## Conclusion

Individual differences in patterns of thoughts, feelings, and behavior tend to stabilize over development. Along with increases in phenotypic stability, genetic and environmental influences both increase in stability with age. Near age 30, genetic stability approaches unity, and true environmental stability slowly increases across the majority of the life span to reach similar levels of stability in old age. The genetic and environmental decomposition of phenotypic stability is likely the most surprising and informative finding of the present study. Genetic influences exert a relatively constant influence on stability across the life span and fully explain phenotypic stability at young ages. Environmental contributions to stability, on the other hand, are almost nonexistent in early childhood, but by midlife, the environment contributes only slightly less to phenotypic stability than do genetic influences. This result indicates that the trend of increasing phenotypic stability can largely be explained by increasing environmental contributions. Previous research has identified instances when the influence of the environment on personality depends on genotype and instances of genetic predispositions leading to nonrandom exposure to the environment. A crucial next step for personality theory and research will be to document how the effects of such gene-environment interplay carry forward in time.

## References

References marked with an asterisk indicate studies included in the meta-analysis.

- Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18*. Burlington: University of Vermont, Department of Psychiatry.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York, NY: Holt, Rinehart, & Winston.
- Andershed, H., Kerr, M., Stattin, H., & Levander, S. (2002). Psychopathic traits in non-referred youths: Initial test of a new assessment tool. In E. Blaauw & L. Sheridan (Eds.), *Psychopaths: Current international perspectives* (pp. 131-158). The Hague, the Netherlands: Elsevier.
- Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*, 469-480. doi:10.1037/0003-066X.55.5.469
- Avinun, R., & Knafo, A. (2014). Parenting as a reaction evoked by children's genotype: A meta-analysis of children-as-twins studies. *Personality and Social Psychology Review*, *18*, 87-102. doi:10.1177/1088868313498308
- Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2006). Gene-environment interaction of the dopamine D4 receptor (DRD4) and observed maternal insensitivity predicting externalizing behavior in preschoolers. *Developmental Psychobiology*, *48*, 406-409. doi:10.1002/dev.20152
- Barenbaum, N. B., & Winter, D. G. (2008). History of modern personality theory and research. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality* (pp. 3-28). New York, NY: Guilford Press.
- Bayley, N. (1969). *Manual for the Bayley Scales of Infant Development*. New York, NY: Psychological Corporation.
- Bell, R. Q. (1953). Convergence: An accelerated longitudinal approach. *Child Development*, *24*, 145-152.
- Belsky, J., & Beaver, K. M. (2011). Cumulative-genetic plasticity, parenting and adolescent self-regulation. *Journal of Child Psychology and Psychiatry*, *52*, 619-626. doi:10.1111/j.1469-7610.2010.02327.x
- Bleidorn, W. (2012). Hitting the road to adulthood: Short-term personality development during a major life transition. *Personality and Social Psychology Bulletin*, *38*, 1594-1608. doi:10.1177/0146167212456707
- Bleidorn, W., Kilmstra, T. A., Denissen, J. J. A., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2013). Personality maturation around the world: A cross-cultural examination of social investment theory. *Psychological Science*, *24*, 2530-2540. doi:10.1177/0956797613498396
- \*Blonigen, D. M., Carlson, M. D., Hicks, B. M., Krueger, R. F., & Iacono, W. G. (2008). Stability and change in personality traits from late adolescence to early adulthood: A longitudinal twin study. *Journal of Personality*, *76*, 229-266. doi:10.1111/j.1467-6494.2007.00485.x
- \*Blonigen, D. M., Hicks, B. M., Krueger, R. F., Patrick, C. J., & Iacono, W. G. (2006). Continuity and change in psychopathic traits as measured via normal-range personality: A longitudinal-biometric study. *Journal of Abnormal Psychology*, *115*, 85-95. doi:10.1037/0021-843X.115.1.85
- Bogg, T., Finn, P. R., & Monsey, K. E. (2012). A year in the college life: Evidence for the social investment hypothesis via trait self-control and alcohol consumption. *Journal of Research in Personality*, *46*, 694-699. doi:10.1016/j.jrp.2012.08.004
- Bouchard, T. J., Jr., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics*, *31*, 243-273. doi:10.1023/A:1012294324713
- Bowlby, J. (1973). *Attachment and loss: Vol. II. Separation: Anxiety and anger*. New York, NY: Basic Books.
- \*Bratko, D., & Butkovic, A. (2007). Stability of genetic and environmental effects from adolescence to young adulthood: Results of Croatian longitudinal twin study of personality. *Twin Research and Human Genetics*, *10*, 151-157. doi:10.1375/twin.10.1.151
- Briley, D. A., & Tucker-Drob, E. M. (2012). Broad bandwidth or high fidelity? Evidence from the structure of genetic and environmental effects on the facets of the five factor model. *Behavior Genetics*, *42*, 743-763. doi:10.1007/s10519-012-9548-8
- Burt, S. A. (2008). Genes and popularity: Evidence of an evocative gene-environment correlation. *Psychological Science*, *19*, 112-113. doi:10.1111/j.1467-9280.2008.02055.x
- Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Erlbaum.
- Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*, 319-333. doi:10.1037/0022-3514.67.2.319
- Caspi, A., Harrington, H., Milne, B., Amell, J. W., Theodore, R. F., & Moffitt, T. E. (2003). Children's behavioral styles at age 3 are linked to their adult personality traits at age 26. *Journal of Personality*, *71*, 495-514. doi:10.1111/1467-6494.7104001
- Caspi, A., & Roberts, B. W. (2001). Personality development across the life course: The argument for change and continuity. *Psychological Inquiry*, *12*, 49-66. doi:10.1207/S15327965PLI1202\_01
- Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, *56*, 453-484. doi:10.1146/annurev.psych.55.090902.141913
- Cheung, M. W. L. (2008). A model for integrating fixed-, random-, and mixed-effects meta-analyses into structural equation modeling. *Psychological Methods*, *13*, 182-202. doi:10.1037/a0013163

- Church, A. T. (1994). Relating the Tellegen and five-factor models of personality structure. *Journal of Personality and Social Psychology, 67*, 898–909. doi:10.1037/0022-3514.67.5.898
- Cloninger, C. R. (1998). The genetics and psychobiology of the seven-factor model of personality. In K. R. Silk (Ed.), *Biology of personality disorders* (pp. 63–92). Washington, DC: American Psychiatric Association.
- Conley, J. J. (1984). The hierarchy of consistency: A review and model of longitudinal findings on adult individual differences in intelligence, personality and self-opinion. *Personality and Individual Differences, 5*, 11–25. doi:10.1016/0191-8869(84)90133-8
- Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: Sage.
- Costa, P. T., Jr., & McCrae, R. R. (1989). *NEO Personality Inventory: Manual Form S and Form R*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., Jr., & McCrae, R. R. (2006). Age changes in personality and their origins: Comment on Roberts, Walton, and Viechtbauer (2006). *Psychological Bulletin, 132*, 26–28. doi:10.1037/0033-2909.132.1.26
- Dall, S. R. X., Houston, A. I., & McNamara, J. M. (2004). The behavioural ecology of personality: Consistent individual differences from an adaptive perspective. *Ecology Letters, 7*, 734–739. doi:10.1111/j.1461-0248.2004.00618.x
- Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., & Damasio, A. R. (1994, May 20). The return of Phineas Gage: Clues about the brain from the skull of a famous patient. *Science, 264*, 1102–1105. doi:10.1126/science.8178168
- \*De Fruyt, F., Bartels, M., Van Leeuwen, K. G., De Clercq, B., Decuyper, M., & Mervielde, I. (2006). Five types of personality continuity in childhood and adolescence. *Journal of Personality and Social Psychology, 91*, 538–552. doi:10.1037/0022-3514.91.3.538
- de Moor, M. H. M., Costa, P. T., Terracciano, A., Krueger, R. F., de Geus, E. J. C., Toshiko, T., . . . Boomsma, D. I. (2012). Meta-analysis of genome-wide association studies for personality. *Molecular Psychiatry, 17*, 337–349. doi:10.1038/mp.2010.128
- Dick, D. M. (2011). Gene-environment interaction in psychological traits and disorders. *Annual Review of Clinical Psychology, 7*, 393–409. doi:10.1146/annurev-clinpsy-032210-104518
- Dick, D. M., Agrawal, A., Schuckit, M. A., Bierut, L., Hinrichs, A., Fox, L., . . . Begleiter, H. (2006). Marital status, alcohol dependence, and GABRA2: Evidence for gene-environment correlation and interaction. *Journal of Studies on Alcohol, 67*, 185–194.
- Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*, 346–369. doi:10.1037/0033-295X.108.2.346
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417–440. doi:10.1146/annurev.ps.41.020190.002221
- Donnellan, M. B., Conger, R. D., & Burzette, B. G. (2005). Criterion-related validity, self-other agreement, and longitudinal analyses for the Iowa Personality Questionnaire: A short alternative to the MPQ. *Journal of Research in Personality, 39*, 458–485. doi:10.1016/j.jrp.2004.05.004
- Donnellan, M. B., Conger, R. D., & Burzette, R. G. (2007). Personality development from late adolescence to young adulthood: Differential stability, normative maturity, and evidence for the maturity-stability hypothesis. *Journal of Personality, 75*, 237–264. doi:10.1111/j.1467-6494.2007.00438.x
- \*Ericson, M., Tuvblad, C., Raine, A., Young-Wolff, K., & Baker, L. A. (2011). Heritability and longitudinal stability of schizotypal traits during adolescence. *Behavior Genetics, 41*, 499–511. doi:10.1007/s10519-010-9401-x
- Erikson, E. H. (1963). *Childhood and society* (2nd ed.). New York, NY: Norton. (Original work published 1950)
- Eysenck, H. J. (1992). Four ways five factors are not basic. *Personality and Individual Differences, 13*, 667–673. doi:10.1016/0191-8869(92)90237-J
- Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. doi:10.1007/978-1-4613-2413-3
- Eysenck, H. J., & Eysenck, S. B. (1975). *Manual of the Eysenck Personality Questionnaire*. San Diego, CA: Educational and Industrial Testing Service.
- Ferguson, C. J. (2010). A meta-analysis of normal and disordered personality across the life span. *Journal of Personality and Social Psychology, 98*, 659–667. doi:10.1037/a0018770
- \*Forsman, M., Lichtenstein, P., Andershed, H., & Larsson, H. (2008). Genetic effects explain the stability of psychopathic personality from mid- to late adolescence. *Journal of Abnormal Psychology, 117*, 606–617. doi:10.1037/0021-843X.117.3.606
- Fraley, R. C., & Roberts, B. W. (2005). Patterns of continuity: A dynamic model for conceptualizing the stability of individual differences in psychological constructs across the life course. *Psychological Review, 112*, 60–74. doi:10.1037/0033-295X.112.1.60
- Freud, S. (1959). Character and anal eroticism. In J. Strachey (Ed. & Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 9, pp. 169–175). London, England: Hogarth Press. (Original work published 1908)
- Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception*. San Diego, CA: Academic Press.
- \*Gagne, J. R., & Goldsmith, H. H. (2011). A longitudinal analysis of anger and inhibitory control in twins from 12 to 36 months of age. *Developmental Science, 14*, 112–124. doi:10.1111/j.1467-7687.2010.00969.x
- \*Ganiban, J. M., Saudino, K. J., Ulbricht, J., Neiderhiser, J. M., & Reiss, D. (2008). Stability and change in temperament during adolescence. *Journal of Personality and Social Psychology, 95*, 222–236. doi:10.1037/0022-3514.95.1.222
- Ganiban, J. M., Ulbricht, J., Saudino, K. J., Reiss, D., & Neiderhiser, J. M. (2011). Understanding child-based effects on parenting: Temperament as a moderator of genetic and environmental contributions to parenting. *Developmental Psychology, 47*, 676–692. doi:10.1037/a0021812
- \*Gillespie, N. A., Evans, D. E., Wright, M. M., & Martin, N. G. (2004). Genetic simplex modeling of Eysenck's dimensions of personality in a sample of young Australian twins. *Twin Research, 7*, 637–648. doi:10.1375/1369052042663814
- Goldsmith, H. H. (1996). Studying temperament via construction of the Toddler Behavior Assessment Questionnaire. *Child Development, 67*, 218–235. doi:10.2307/1131697
- Heath, A. C., Berg, K., Eaves, L. J., Solaas, M. H., Corey, L. A., Sundet, J., . . . Nance, W. E. (1985, April 25). Education policy and the heritability of educational attainment. *Nature, 314*, 734–736. doi:10.1038/314734a0
- Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3*, 486–504. doi:10.1037/1082-989X.3.4.486
- \*Hopwood, C. J., Donnellan, M. B., Blonigen, D. M., Krueger, R. F., McGue, M., Iacono, W. G., & Burt, S. A. (2011). Genetic and environmental influences on personality trait stability and growth during the transition to adulthood: A three-wave longitudinal study. *Journal of Personality and Social Psychology, 100*, 545–556. doi:10.1037/a0022409
- Jackson, J. J., Hill, P. L., Payne, B. R., Roberts, B. W., & Stine-Morrow, E. L. (2012). Can an old dog learn (and want to experience) new tricks? Cognitive training increases openness to experience in older adults. *Psychology and Aging, 27*, 286–292. doi:10.1037/a0025918

- Jackson, J. J., Thoemmes, F., Jonkman, K., Lüdtke, O., & Trautwein, U. (2012). Military training and personality trait development: Does the military make the man, or does the man make the military? *Psychological Science*, *23*, 270–277. doi:10.1177/0956797611423545
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality* (pp. 114–158). New York, NY: Guilford Press.
- Johnson, W. (2007). Genetic and environmental influences on behavior: Capturing all the interplay. *Psychological Review*, *114*, 423–440. doi:10.1037/0033-295X.114.2.423
- \*Johnson, W., McGue, M., & Krueger, R. F. (2005). Personality stability in late adulthood: A behavior genetic analysis. *Journal of Personality*, *73*, 523–552. doi:10.1111/j.1467-6494.2005.00319.x
- Johnson, W., Penke, L., & Spinath, F. M. (2011). Heritability in the era of molecular genetics: Some thoughts for understanding genetic influences on behavioral traits. *European Journal of Personality*, *25*, 254–266. doi:10.1002/per.836
- Kandler, C. (2012). Nature and nurture in personality development: The case of neuroticism and extraversion. *Current Directions in Psychological Science*, *21*, 290–296. doi:10.1177/0963721412452557
- Kandler, C., Bleidorn, W., Riemann, R., Angleitner, A., & Spinath, F. M. (2012). Life events as environmental states and genetic traits and the role of personality: A longitudinal twin study. *Behavior Genetics*, *42*, 57–72. doi:10.1007/s10519-011-9491-0
- \*Kandler, C., Bleidorn, W., Riemann, R., Spinath, F. M., Thiel, W., & Angleitner, A. (2010). Sources of cumulative continuity in personality: A longitudinal multiple-rater twin study. *Journal of Personality and Social Psychology*, *98*, 995–1008. doi:10.1037/a0019558
- Kaufman, J., Yang, B.-Z., Douglas-Palumberi, H., Grasso, D., Lipschitz, D., Houshyar, S., . . . Gelernter, J. (2006). Brain-derived neurotrophic factor–5-HTTLPR gene interactions and environmental modifiers of depression in children. *Biological Psychiatry*, *59*, 673–680. doi:10.1016/j.biopsych.2005.10.026
- Keller, M. C., Coventry, W. L., Heath, A. C., & Martin, N. G. (2005). Widespread evidence for non-additive genetic variation in Cloninger's and Eysenck's personality dimensions using a twin plus sibling design. *Behavior Genetics*, *35*, 707–721. doi:10.1007/s10519-005-6041-7
- Kendler, K. S., & Baker, J. H. (2007). Genetic influences on measures of the environment: A systematic review. *Psychological Medicine*, *37*, 615–626. doi:10.1017/S0033291706009524
- Klahr, A. M., & Burt, S. A. (2014). Elucidating the etiology of individual differences in parenting: A meta-analysis of behavioral genetic research. *Psychological Bulletin*, *140*, 544–586. doi:10.1037/a0034205
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.
- Krasner, M. S., Epstein, R. M., Beckman, H., Suchman, A. L., Chapman, B., Mooney, C. J., & Quill, T. E. (2009). Association of an educational program in mindful communication with burnout, empathy, and attitudes among primary care physicians. *JAMA: Journal of the American Medical Association*, *302*, 1284–1293. doi:10.1001/jama.2009.1384
- Krueger, R. F., & Johnson, W. (2008). Behavior genetics and personality: A new look at the integration of nature and nurture. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality* (pp. 287–310). New York, NY: Guilford Press.
- Krueger, R. F., Johnson, W., & Kling, K. C. (2006). Behavior genetics and personality development. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 81–109). Mahwah, NJ: Erlbaum.
- Krueger, R. F., South, S., Johnson, W., & Iacono, W. (2008). The heritability of personality is not always 50%: Gene-environment interactions and correlations between personality and parenting. *Journal of Personality*, *76*, 1485–1522. doi:10.1111/j.1467-6494.2008.00529.x
- \*Kupper, N., Boomsma, D. I., de Geus, E. J. C., Denollet, J., & Willemsen, G. (2011). Nine-year stability of Type D personality: Contributions of genes and environment. *Psychosomatic Medicine*, *73*, 75–82. doi:10.1097/PSY.0b013e3181fdce54
- Kupper, N., & Denollet, J. (2007). Type D personality as a prognostic factor in heart disease: Assessment and mediating mechanisms. *Journal of Personality Assessment*, *89*, 265–276. doi:10.1080/00223890701629797
- Lindová, J., Příplatová, L., & Flegr, J. (2012). Higher extraversion and lower conscientiousness in humans infected with *Toxoplasma*. *European Journal of Personality*, *26*, 285–291. doi:10.1002/per.838
- Loehlin, J. C. (1992). *Genes and environment in personality development*. Thousand Oaks, CA: Sage.
- Lucas, R. E., & Donnellan, M. B. (2011). Personality development across the life span: Longitudinal analyses with a national sample from Germany. *Journal of Personality and Social Psychology*, *101*, 847–861. doi:10.1037/a0024298
- Lucht, M., Barnow, S., Schroeder, W., Grabe, H. J., Finckh, U., John, U., . . . Herrmann, F. H. (2006). Negative perceived parenting is associated with dopamine D<sub>2</sub> receptor exon 8 and GABA(A) alpha 6 receptor variants: An explorative study. *American Journal of Medical Genetics: B. Neuropsychiatric Genetics*, *141B*, 167–172. doi:10.1002/ajmg.b.30255
- Luijk, M. P. C. M., Roisman, G. I., Haltigan, J. D., Tiemeier, H., Booth-LaForce, C., van IJzendoorn, M. H., . . . Bakermans-Kranenburg, M. J. (2011). Dopaminergic, serotonergic, and oxytonergic candidate genes associated with infant attachment security and disorganization? In search of main and interaction effects. *Journal of Child Psychology and Psychiatry*, *52*, 1295–1307. doi:10.1111/j.1469-7610.2011.02440.x
- Maher, B. (2008, November 5). Personal genomes: The case of the missing heritability. *Nature*, *456*, 18–21. doi:10.1038/456018a
- McAdams, D. P., & Olson, B. D. (2010). Personality development: Continuity and change over the life course. *Annual Review of Psychology*, *61*, 517–542. doi:10.1146/annurev.psych.093008.100507
- McAdams, T. A., Gregory, A. M., & Eley, T. C. (2013). Genes of experience: Explaining the heritability of putative environmental variables through their association with behavioral and emotional traits. *Behavior Genetics*, *43*, 314–328. doi:10.1007/s10519-013-9591-0
- McCartney, K., Harris, M. J., & Bernieri, F. (1990). Growing up and growing apart: A developmental meta-analysis of twin studies. *Psychological Bulletin*, *107*, 226–237. doi:10.1037/0033-2909.107.2.226
- McClintock, M. K., & Herdt, G. (1996). Rethinking puberty: The development of sexual attraction. *Current Directions in Psychological Science*, *5*, 178–183. doi:10.1111/1467-8721.ep11512422
- McCrae, R. R., & Costa, P. T., Jr. (1994). The stability of personality: Observations and evaluations. *Current Directions in Psychological Science*, *3*, 173–175. doi:10.1111/1467-8721.ep10770693
- McCrae, R. R., & Costa, P. T., Jr. (2006). Cross-cultural perspectives on adult personality trait development. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 129–146). Mahwah, NJ: Erlbaum.
- McCrae, R. R., & Costa, P. T., Jr. (2008). The five-factor theory of personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality* (pp. 159–181). New York, NY: Guilford Press.
- McCrae, R. R., Costa, P. T., Jr., de Lima, M., Simões, A., Ostendorf, F., Angleitner, A., . . . Piedmont, R. L. (1999). Age differences in personality across the adult life span: Parallels in five cultures. *Developmental Psychology*, *35*, 466–477. doi:10.1037/0012-1649.35.2.466
- McCrae, R. R., Costa, P. T., Jr., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M. D., . . . Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology*, *78*, 173–186. doi:10.1037/0022-3514.78.1.173



- \*McGue, M., Bacon, S., & Lykken, D. T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental Psychology, 29*, 96–109. doi:10.1037/0012-1649.29.1.96
- Measelle, J. R., John, O. P., Ablow, J. C., Cowan, P. A., & Cowan, C. P. (2005). Can children provide coherent, stable, and valid self-reports on the Big Five dimensions? A longitudinal study from ages 5 to 7. *Journal of Personality and Social Psychology, 89*, 90–106. doi:10.1037/0022-3514.89.1.90
- 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 Eighth European Conference on Personality Psychology* (pp. 107–127). Tilburg, the Netherlands: Tilburg University Press.
- Mroczek, D. K., & Little, T. D. (2006). Theory and research in personality development at the beginning of the 21st century. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 3–10). Mahwah, NJ: Erlbaum.
- Murray, H. A. (1938). *Explorations in personality: A clinical and experimental study of fifty men of college age*. New York, NY: Oxford University Press.
- Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.
- Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. doi:10.1007/978-94-015-8018-2
- \*Niv, S., Tuvblad, C., Raine, A., Wang, P., & Baker, L. A. (2012). Heritability and longitudinal stability of impulsivity in adolescence. *Behavior Genetics, 42*, 378–392. doi:10.1007/s10519-011-9518-6
- Ones, D. S., & Viswesvaran, C. (2001). Personality at work: Criterion-focused occupational personality scales used in personnel selection. In B. W. Roberts & R. Hogan (Eds.), *Personality psychology in the workplace* (pp. 63–92). doi:10.1037/10434-003
- Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology, 51*, 768–774. doi:10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1
- \*Pedersen, N. L., & Reynolds, C. A. (1998). Stability and change in adult personality: Genetic and environmental components. *European Journal of Personality, 12*, 365–386. doi:10.1002/(SICI)1099-0984(199809)12:5<365::AID-PER335>3.0.CO;2-N
- Penke, L., Denissen, J. J. A., & Miller, G. F. (2007). The evolutionary genetics of personality. *European Journal of Personality, 21*, 549–587. doi:10.1002/per.629
- Plomin, R. (2013). Child development and molecular genetics: 14 years later. *Child Development, 84*, 104–120. doi:10.1111/j.1467-8624.2012.01757.x
- Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioral and Brain Sciences, 10*, 1–16. doi:10.1017/S0140525X00055941
- Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84*, 309–322. doi:10.1037/0033-2909.84.2.309
- Plomin, R., Reiss, D., Hetherington, E. M., & Howe, G. W. (1994). Nature and nurture: Genetic contributions to measures of the family environment. *Developmental Psychology, 30*, 32–43. doi:10.1037/0012-1649.30.1.32
- Purcell, S. (2002). Variance components models for gene–environment interaction in twin analysis. *Twin Research and Human Genetics, 5*, 554–571. doi:10.1375/twin.5.6.554
- Raine, A., & Baker, L. A. (2003). *The SPQ-C: A brief child version of the Schizotypal Personality Questionnaire*. Unpublished questionnaire, Department of Psychology, University of Southern California, Los Angeles, California.
- \*Read, S., Vogler, G. P., Pedersen, N. L., & Johansson, B. (2006). Stability and change in genetic and environmental components of personality in old age. *Personality and Individual Differences, 40*, 1637–1647. doi:10.1016/j.paid.2006.01.004
- Rettew, D. C., Rebollo-Mesa, I., Hudziak, J. J., Willemsen, G., & Boomsma, D. I. (2008). Non-additive and additive genetic effects on extraversion in 3314 Dutch adolescent twins and their parents. *Behavior Genetics, 38*, 223–233. doi:10.1007/s10519-008-9192-5
- Rhemtulla, M., & Tucker-Drob, E. M. (2012). Gene-by-socioeconomic status interaction on school readiness. *Behavior Genetics, 42*, 549–558. doi:10.1007/s10519-012-9527-0
- Riemann, R., Angleitner, A., & Strelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self- and peer report NEO-FFI scales. *Journal of Personality, 65*, 449–475. doi:10.1111/j.1467-6494.1997.tb00324.x
- Riemann, R., Kandler, C., & Bleidorn, W. (2012). Behavioral genetic analyses of parent twin relationship quality. *Personality and Individual Differences, 53*, 398–404. doi:10.1016/j.paid.2012.02.022
- \*Rietveld, M. J. H., Hudziak, J. J., Bartels, M., van Beijsterveldt, C. E. M., & Boomsma, D. I. (2004). Heritability of attention problems in children: Longitudinal results from a study of twins, age 3 to 12. *Journal of Child Psychology and Psychiatry, 45*, 577–588. doi:10.1111/j.1469-7610.2004.00247.x
- Roberts, B. W. (2009). Back to the future: *Personality and Assessment and personality development*. *Journal of Research in Personality, 43*, 137–145. doi:10.1016/j.jrp.2008.12.015
- 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. doi:10.1037/0022-3514.81.4.670
- 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. doi:10.1037/0033-2909.126.1.3
- Roberts, B. W., & Jackson, J. J. (2008). Sociogenomic personality psychology. *Journal of Personality, 76*, 1523–1544. doi:10.1111/j.1467-6494.2008.00530.x
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*, 1–25. doi:10.1037/0033-2909.132.1.1
- Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 11–41). Mahwah, NJ: Erlbaum.
- Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality* (pp. 375–398). New York, NY: Guilford Press.
- Roberts, B. W., Wood, D., & Smith, J. L. (2005). Evaluating five factor theory and social investment perspectives on personality trait development. *Journal of Research in Personality, 39*, 166–184. doi:10.1016/j.jrp.2004.08.002
- Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development, 52*, 569–578. doi:10.2307/1129176
- Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Child's Behavior Questionnaire. *Child Development, 72*, 1394–1408. doi:10.1111/1467-8624.00355
- \*Saudino, K. J. (2012). Sources of continuity and change in activity level in early childhood. *Child Development, 83*, 266–281. doi:10.1111/j.1467-8624.2011.01680.x
- \*Saudino, K. J., & Cherny, S. S. (2001). Sources of continuity and change



- in observed temperament. In R. N. Emde & J. K. Hewitt (Eds.), *Infancy to early childhood: Genetic and environmental influences on developmental change* (pp. 89–110). New York, NY: Oxford University Press.
- Saudino, K. J., & Plomin, R. (1997). Cognitive and temperamental mediators of genetic contributions to the home environment during infancy. *Merrill-Palmer Quarterly*, *43*, 1–23.
- Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development*, *54*, 424–435.
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*, 350–353. doi:10.1037/1040-3590.8.4.350
- Sheese, B. E., Voelker, P. M., Rothbart, M. K., & Posner, M. I. (2007). Parenting quality interacts with genetic variation in dopamine receptor D4 to influence temperament in early childhood. *Development and Psychopathology*, *19*, 1039–1046. doi:10.1017/S0954579407000521
- Sheldon, W. H., Stevens, S. S., & Tucker, W. B. (1940). *The varieties of human physique: An introduction to constitutional psychology*. New York, NY: Harper.
- Shiner, R. L., Masten, A. S., & Tellegen, A. (2002). A developmental perspective on personality in emerging adulthood: Childhood antecedents and concurrent adaptation. *Journal of Personality and Social Psychology*, *83*, 1165–1177. doi:10.1037/0022-3514.83.5.1165
- Silverman, W. K., Fleisig, W., Rabian, B., & Peterson, R. A. (1991). Childhood Anxiety Sensitivity Index. *Journal of Clinical Child Psychology*, *20*, 162–168. doi:10.1207/s15374424jccp2002\_7
- Smith, H. J., Sheikh, H. I., Dyson, M. W., Olino, T. M., Lupton, R. S., Durbin, C. E., . . . Klein, D. N. (2012). Parenting and child DRD4 genotype interact to predict children's early emerging effortful control. *Child Development*, *83*, 1932–1944. doi:10.1111/j.1467-8624.2012.01818.x
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of Big Five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, *94*, 718–737. doi:10.1037/0022-3514.94.4.718
- \*Spengler, M., Gottschling, J., & Spinath, F. M. (2012). Personality in childhood: A longitudinal behavior genetic approach. *Personality and Individual Differences*, *53*, 411–416. doi:10.1016/j.paid.2012.01.019
- Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology*, *84*, 1041–1053. doi:10.1037/0022-3514.84.5.1041
- Sturaro, C., Denissen, J. J. A., van Aken, M. A. G., & Asendorpf, J. B. (2008). Person-environment transactions during emerging adulthood: The interplay between personality characteristics and social relationships. *European Psychologist*, *13*, 1–11. doi:10.1027/1016-9040.13.1.1
- Tabery, J. (2007). Biometric and developmental gene–environment interactions: Looking back, moving forward. *Development and Psychopathology*, *19*, 961–976. doi:10.1017/S0954579407000478
- Tackett, J. L., Slobodskaya, H. R., Mar, R. A., Deal, J., Halverson, C. F., Jr., Baker, S. R., . . . Besevegis, E. (2012). The hierarchical structure of childhood personality in five countries: Continuity from early childhood to early adolescence. *Journal of Personality*, *80*, 847–879. doi:10.1111/j.1467-6494.2011.00748.x
- \*Takahashi, Y., Yamagata, S., Kijima, N., Shigemasu, K., Ono, Y., & Ando, J. (2007). Continuity and change in behavioral inhibition and activation systems: A longitudinal behavior genetic study. *Personality and Individual Differences*, *43*, 1616–1625. doi:10.1016/j.paid.2007.04.030
- Tang, T. Z., DeRubeis, R. J., Hollon, S. D., Amsterdam, J., Shelton, R., & Schalet, B. (2009). Personality change during depression treatment: A placebo-controlled trial. *Archives of General Psychiatry*, *66*, 1322–1330. doi:10.1001/archgenpsychiatry.2009.166
- Tellegen, A., & Waller, N. G. (2008). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *Sage handbook of personality theory and testing, Vol. 2: Personality measurement and assessment* (pp. 254–285). London, England: Sage.
- Terracciano, A., Costa, P. T., Jr., & McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin*, *32*, 999–1009. doi:10.1177/0146167206288599
- Terracciano, A., McCrae, R. R., & Costa, P. T., Jr. (2010). Intra-individual change in personality stability and age. *Journal of Research in Personality*, *44*, 31–37. doi:10.1016/j.jrp.2009.09.006
- Terracciano, A., Sanna, S., Uda, M., Deiana, B., Usala, G., Busonero, F., . . . Costa, P. T., Jr. (2010). Genome-wide association scan for five major dimensions of personality. *Molecular Psychiatry*, *15*, 647–656. doi:10.1038/mp.2008.113
- Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science*, *9*, 160–164. doi:10.1111/1467-8721.00084
- Turkheimer, E., & Waldron, M. (2000). Nonshared environment: A theoretical, methodological, and quantitative review. *Psychological Bulletin*, *126*, 78–108. doi:10.1037/0033-2909.126.1.78
- Verweij, K. J. H., Zietsch, B. P., Medland, S. E., Gordon, S. D., Benjamin, B., Nyholt, D. R., . . . Wray, N. R. (2010). A genome-wide association study of Cloninger's temperament scales: Implications for the evolutionary genetics of personality. *Biological Psychology*, *85*, 306–317. doi:10.1016/j.biopsycho.2010.07.018
- \*Viken, R. J., Rose, R. J., Kaprio, J., & Koskenvuo, M. (1994). A developmental genetic analysis of adult personality: Extraversion and neuroticism from 18 to 59 years of age. *Journal of Personality and Social Psychology*, *66*, 722–730. doi:10.1037/0022-3514.66.4.722
- Vinkhuyzen, A. A. E., Pedersen, N. L., Yang, J., Lee, S. H., Magnusson, P. K. E., Iacono, W. G., . . . Wray, N. R. (2012). Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. *Translational Psychiatry*, *2*, Article e102. doi:10.1038/tp.2012.27
- Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*, *17*, 228–243. doi:10.1037/a0027127
- Webb, B. T., Guo, A.-Y., Maher, B. S., Zhao, Z., van den Oord, E. J. V., Kendler, K. S., . . . Hettema, J. M. (2012). Meta-analyses of genome-wide linkage scans of anxiety-related phenotypes. *European Journal of Human Genetics*, *20*, 1078–1084. doi:10.1038/ejhg.2012.47
- Wolf, M., & Weissing, F. J. (2012). Animal personalities: Consequences for ecology and evolution. *Trends in Ecology and Evolution*, *27*, 452–461. doi:10.1016/j.tree.2012.05.001
- \*Zavos, H. M. S., Gregory, A. M., & Eley, T. C. (2012). Longitudinal genetic analysis of anxiety sensitivity. *Developmental Psychology*, *48*, 204–212. doi:10.1037/a0024996
- Zhao, H.-K., Chang, C.-W., Geng, N., Gao, L., Wang, J., Wang, X., . . . Gao, G.-D. (2012). Associations between personality changes and nucleus accumbens ablation in opioid addicts. *Acta Pharmacologica Sinica*, *33*, 588–593. doi:10.1038/aps.2012.10
- Zimmermann, J., & Neyer, F. J. (2013). Do we become a different person when hitting the road? Personality development of sojourners. *Journal of Personality and Social Psychology*, *105*, 515–530. doi:10.1037/a0033019

Received July 12, 2013

Revision received April 25, 2014

Accepted April 28, 2014 ■