The Dissertation Committee for Daniel Andrew Briley Certifies that this is the approved version of the following dissertation:

The Structure, Development, and Correlates of Individual Differences in Human Personality

Committee:

Elliot M. Tucker-Drob, Supervisor

K. Paige Harden

Samuel D. Gosling

David M. Buss

Joseph E. Potter

Robert Crosnoe

The Structure, Development, and Correlates of Individual Differences in Human Personality

by

Daniel Andrew Briley, B.A.; M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2015

The Structure, Development, and Correlates of Individual Differences in Human Personality

Daniel Andrew Briley, Ph.D.

The University of Texas at Austin, 2015

Supervisor: Elliot M. Tucker-Drob

Human personality possesses a complex psychometric structure that changes with maturation and predicts a number of important life outcomes. The current dissertation presents four studies that track the structure, development, and correlates of personality, primarily through a behavioral genetic lens. Study 1 finds that the genetic and environmental structure of personality domains residing over more narrow facets is more complex than can be accounted for by simple, broad domains. Study 2 presents meta-analytic evidence for the differential stability of personality traits across the lifespan, as well as genetic and environmental contributions to stability. Personality traits become more stable with age, and this trend is primarily driven by increasing environmental sources of stability. Study 3 details a transactional model of parental educational expectations, child academic behaviors, and child academic achievement across early development. Finally, Study 4 presents associations between state-level aggregates of personality and the level, timing, and context of fertility across the United States. Together, these studies indicate the importance of understanding personality for individual-level and population-level processes, as well as, the complexities of this goal due to the dynamic nature of gene-environment dependencies that undergird personality development.

Table of Contents

List of Tables

List of Figures

INTRODUCTION

Personality refers to differential patterns of cognition, emotion, and behavior that are relatively consistent across time and context. What processes shape the formation of these differences, and what are the potential consequences of these differences for important life outcomes? It is clear that individual differences in personality result from both genetic variation across individuals and variation in life experiences, along with complex interactions between genetic predispositions and environmental circumstances (Bouchard & Loehlin, 2001; Krueger & Johnson, 2008). Further, individual differences in personality have important real-world consequences. In a number of high quality longitudinal studies and comprehensive meta-analyses, individual differences in personality have been found to be predictive of longevity, mental health, academic achievement, economic prosperity, and divorce (Caspi et al., 2014; Cuijpers et al., 2010; Deary, Weiss, & Batty, 2010; Moffitt et al., 2011; Poropat, 2009; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). Finally, individual differences in human personality fall along a number of statistical dimensions residing at multiple levels of generality and specificity within a hierarchy, but the extent to which these dimensions represent a developmentally invariant, etiologically "crisp" psychometric structure is not well known. The current dissertation presents four studies that investigate the structure, development, and correlates of personality, primarily through a behavioral genetic lens.

Study 1 explores the hierarchical structure of the broad Big Five personality domains (i.e. extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience) residing over more narrow personality facets (Briley & Tucker-Drob,

2012). The psychometric structure of within domain facets has previously been assumed to be highly organized around the broad domain (e.g., McCrae & Costa, 2008). This study evaluates this hypothesis by contrasting independent and common pathways behavioral genetic models. Independent pathways models estimate a relatively unconstrained account of the variance that is shared among facets within a domain, whereas the common pathways model imposes a common factor structure to the shared variance. I find that the genetic and environmental factor structure is more complex than can be accounted for by simple, broad domain factors. Moreover, most facets reflect a combination of common and specific genetic and environmental influences.

Building off the foundational meta-analysis by Roberts and DelVecchio (2000) that documents increasing differential stability of personality across the lifespan, Study 2 uses meta-analysis to investigate developmental trends in the longitudinal stabilities of genetic and environmental influences on personality and their contributions to phenotypic stability. I find that genetic stability reaches unity near age 30, and environmental stability, which is completely absent in early childhood, is appreciable by age 30. Despite these similar trends, I find that the rise in phenotypic stability with age is largely attributable to rising environmental contributions to stability.

Study 3 details a transactional process that unfolds across development between parental educational expectations, child general patterns of academic behavior, and child academic achievement. Using a nationally representative sample of young twins, I find evidence for gene-environment correlations in the form of the parenting that a child receives and genetic influences relevant to the child's general academic behaviors. Moreover, using longitudinal cross-lagged path models, I find that child academic behaviors are a driving force underlying child effects on parental educational expectations.

Finally, Study 4 uses a unique blend of methods from demography and individual differences psychology to test the association between regional differences in personality and fertility, a tremendously important life outcome by both sociological and evolutionary accounts. I investigate gender and age group-specific means in the Big Five personality traits by state, and how these state-to-state differences in personality relate to the timing, level, and structure of aggregate fertility behaviors. Results indicate that states with high levels of agreeableness and conscientiousness tend to have more conventional fertility regimes (i.e., earlier, higher, and more traditional structure), whereas states with high levels of neuroticism and openness tend to reflect the opposite pattern (i.e., later, less, and reduced structure).

Together, these studies indicate the importance of understanding personality for individual-level and population-level processes, and underscore the dynamic nature of genetic and environmental processes that undergird personality development.

Chapter 1: Broad Bandwidth or High Fidelity? Evidence from the Structure of Genetic and Environmental Effects on the Facets of the Five Factor Model

Philosophers and psychologists have long debated whether human personality is undergirded by fundamental dimensions, and if so what these dimensions are. As early as circa 300 BC (see Theophrastus, 1870), Tyrtamus of Eresos in his *Characters* presented the fascinating and intriguing proposition that humans take on several qualitatively different personality types, and this line of investigation has continued to the present day in both popular and academic writings. Currently, the prevailing model of personality structure is the Five Factor Model (FFM; John & Srivastava, 1999). This model posits five broad personality domains that characterize and account for the majority of variation in enduring patterns of how individuals typically behave (Digman, 1990). These domains are Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience (Goldberg, 1993). Several more specific facets are subsumed within each domain. For example, it has been proposed that Conscientiousness is composed of the facets Competence, Order, Dutifulness, Achievement Striving, Self-Discipline, and Deliberation (Costa & McCrae, 1992a). While there is wide acknowledgement that meaningful unique variation exists for each of the facets within the Big Five personality domains, it has become commonplace for contemporary research approaches to focus exclusively on the Big Five domains, which by definition, only include variation that is common among facets of a domain. The current project uses a multivariate behavioral genetic approach to clarify the extent to which these broad domains are capable of capturing genetic variation in the more specific facet measures.

Debates about competing hierarchical taxonomies, their existence or preference for different organizational schemes, is very common in the personality literature.

Particularly, a measurement crisis occurred when several competing models (Block, 1995; Costa & McCrae, 1992b, Eysenck, 1992b; Zuckerman, 1992) were put forward as comprising only basic or fundamental factors. The burden of proof was placed on the newest personality theory, the FFM, to demonstrate that Agreeableness and Conscientiousness were not both subcomponents of a higher order factor, such as Psychoticism, and that Openness to Experience was a personality trait and not a measure of culture or some other non-personality construct (Eysenck, 1992a; McCrae, 1994; McCrae & Costa, 1985; Zuckerman, Kuhlman, Joireman, Teta, & Kraft, 1993). However, evidence in support of the FFM has grown to such a level that the traits of Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience have become paradigmatic reference points for personality research (John & Srivastava, 1999, but see also Block, 2010 for continued disagreement).

BROAD BANDWIDTH

More recently, research about personality structure has centered within or above the Big Five in light of what Cronbach and Gleser (1957) termed the bandwidth-fidelity dilemma. This practical dilemma results from the trade-off between using measures that will cover the majority of variation in personality (domain level measurement) or measures that will assess a few specific behavioral tendencies (facet level measurement) more precisely. Ones and Viswesvaran (1996) have argued that for complex behavioral outcomes, equally broad or complex personality traits are likely to provide the most accurate prediction as they can cover the entire range of behavior patterns important to the outcome. For example, Ones, Viswesvaran, and Schmidt (1993) claimed that for the prediction of job performance, a highly complex outcome typically requiring the ability to work in a social environment, maintain high productivity, and avoid loafing or theft, a

compound personality measure (labeled integrity) reflecting high Conscientiousness and Agreeableness and low Neuroticism is necessary to account for the diverse requirements of most jobs. Further evidence of the usefulness of compound traits comes from research on customer service orientation, violence and aggression, stress tolerance, drug and alcohol use, and self-regulation (Fein & Klein, 2011; Ones & Viswesvaran, 2001a,b).

In light of these results and the common finding that the Big Five are not entirely orthogonal, researchers have proposed two higher order factors, sometimes referred to as Stability and Plasticity, as residing hierarchically above the FFM (DeYoung, 2006; Digman, 1997). The "metatraits" of Stability (represented by high Agreeableness, Conscientiousness, and low Neuroticism) and Plasticity (represented by high Extraversion and Openness) have proved to have evidence of criterion validity (Hirsh, DeYoung, & Peterson, 2009). These high-bandwidth factors have been found to be associated with engagement or restraint of general behavioral acts, individual differences in circadian rhythm, externalizing behaviors, and basic values (DeYoung, Hasher, Djikic, Criger, Peterson, 2007; DeYoung, Peterson, Séguin, & Tremblay, 2008; Hirsch et al., 2009; Vecchione, Alessandri, Barbaranelli, & Caprara, 2011). Stability and Plasticity are thought to represent individual differences in the function of serotonin and dopamine, respectively (Hirsch et al., 2009). The use of highly broad traits is supported by the far-reaching biological effects differences in serotonin and dopamine functioning have on behavior. Additionally, Stability and Plasticity have been linked with components of reinforcement sensitivity theory, namely the behavioral activation system and the behavioral inhibition system, as further evidence of the biological function of these individual differences (Mitchell et al., 2007).

Other researchers take this a step farther and postulate a general factor of personality (Musek, 2007; Rushton & Irwing, 2008, 2009, 2011). These researchers view

the endeavor as having intrinsic worth for understanding genetics, human nature, and evolution (Rushton, Bons, & Hur, 2008), but little research has been conducted on the criterion validity of the extreme high-order factor (van der Linden, te Nijenhuis, & Bakker, 2010). Some exceptions to this are the finding that the general factor is positively related to self-esteem and supervisor-rated job performance (Erdle, Irwing, Rushton, Park, 2010; van der Linden et al., 2010). Similar to Stability and Plasticity, the general factor has been linked with reinforcement sensitivity theory in an effort to provide a biological understanding of why individual differences in the construct exist (Erdle & Rushton, 2010). Much research remains as to whether these findings are merely measurement artifacts (Bäckström, Björklund, & Larsson, 2009; Erdle, Gosling, & Potter, 2009; McCrae et al., 2008), are extensively hindering theoretical development (Ferguson, Chamorro-Premuzic, Pickering, & Weiss, 2011), or are better represented by parsimonious blended traits instead of postulating higher-order factors (Ashton, Lee, Goldberg, & de Vries, 2009).

HIGH FIDELITY

Meanwhile, a minority of researchers have moved in the direction of endorsing more fine-grained measurement (DeYoung, Quilty, & Peterson, 2007). Researchers from this group have argued that lower-level traits may lose broad bandwidth, but they gain fidelity, allowing them to predict more specific behavioral patterns (Paunonen, 1998). For example, Paunonen and Ashton (2001a) demonstrated the possible utility of a narrow rather than broad approach to personality measurement. In this study trained raters carefully selected five facet-level personality dimensions to predict forty behavioral outcomes based on rational grounds in addition to the broad measures of the FFM. The narrow measures of personality were able to account for more variance when included

alone in a regression equation than the broad traits, and importantly, the facets were able to add incremental prediction when entered with broad trait level factors. This indicates that some useful information is lost when facets are aggregated to the broad trait level.

There is some recognition that the use of narrow traits can be more effective for making behavioral predictions. Examples of narrow traits outperforming broad personality factors include the prediction of academic achievement (Luciano, Wainwright, Wright, & Martin, 2006; Paunonen & Ashton, 2001b), broad job performance (Ashton, 1998; Christiansen & Robie, 2011) and specific performance criteria (Dudley, Orvis, Lebiecki, Cortina, 2006), antisocial behavior (Corff & Toupin, 2010), prejudice (Ekehammar &Akrami, 2007), drug use (Terracciano, Löckenhoff, Crum, Bienvenu, & Costa, 2008), longevity (Terracciano, Löckenhoff, Zonderman, Ferrucci, Costa, 2008), and extremes in weight (Terracciano et al., 2009). Additionally, there is evidence that cross-cultural results are more generalizable or reliable when using narrow facet-level measurement (Paunonen, Haddock, Forsterling, & Keinonen, 2003), and that facet-level measurement allows different personality constructs to be mapped on to one another with greater clarity (Armstrong & Anthoney, 2009).

BIG FIVE AS THE STATUS QUO MEASUREMENT LEVEL

Despite these movements to conceptualize personality either higher or lower on the trait hierarchy scheme, the majority of studies measure at the Big Five level. For example, Christiansen & Robie (2011) found that of 200 studies that utilized FFM measurement in the Journal of Applied Psychology and Personnel Psychology over roughly eight years, only 10 studies measured narrow traits. For many approaches, this may simply be a function of practicality. As Goldberg (1993) noted, "Because one always loses specific variance as one amalgamates measures, the optimal level of

prediction is completely a function of statistical power and sample size" (p. 181). Obtaining a sample large enough to precisely measure five broad traits is far easier than the sample required for thirty facets. However, Costa and McCrae (1995) have argued for the practical importance of facet-level measurement for two reasons. First, while each of the facets of the FFM was found to load highly on its intended factor in a principal factors analysis, there was also strong evidence of specific variance associated with each facet that is not held in common with the other facets of a factor. In fact, seven facets actually loaded more strongly on their specific factor rather than the common factor. Second, from an applied perspective, analysis of the relationships of facets within a factor can shed important light on the types of treatment or approaches that will be most effective for a given patient. Utilizing narrow personality measures can move analysis closer to actual mechanisms of behavior.

Clearly, research surrounding the proper level of analysis of personality traits is of direct relevance to pragmatic research goals such as discovering replicable findings that are not simply sample specific. Cronbach (1960) suggested that researchers could initially measure at a broadband level, and as specific replicable findings are presented in the literature, high fidelity measures could then shed light on the presumed mechanisms or processes that the construct represents. In general agreement with this sentiment, Caspi and Shiner (2006) have argued that it is "a short-sighted strategy to rely exclusively on measures of broad superfactors" to make reliable judgments about behavior (p. 332). The argument in favor of this claim is simple. If there are differential effects of some of the more nuanced components of a personality trait, then these possibly very important effects may not be present when measured at the trait level. For example, if half of the facets of Openness are positively related to some important outcome and the other half are negatively related to this outcome, it is conceivable that there will be no measurable

relationship between Openness and the outcome, but this does not mean Openness is not an essential predictor of the outcome. Cronbach (1960, p. 604) points out that although beginning a research program using broad bandwidth measures (such as Openness) may be sensible, lack of relation between a broadly measured trait and an outcome of interest may not necessarily indicate a lack of relation between a more specific aspect of that trait (such as the Openness facet termed Intellectualism) and the outcome of interest. Rather, because broad bandwidth measures are imprecise conglomerates of multiple, potentially distinct subtraits, results of broad bandwidth measures should be taken as a first step in further exploration that includes greater fidelity. A less extreme example of this potential error might be a case in which only one facet of Openness relates to an outcome, and the remaining facets are completely unrelated to the outcome, resulting in a very weak, if even detectable, relation between Openness and the outcome. In both hypothetical scenarios, finer grained analyses would be warranted. Interestingly, rather than moving in this direction, contemporary personality research continues to rely heavily on broad levels of measurement and analysis.

BEHAVIOR GENETICS AS A METHOD TO SHED LIGHT ON STRUCTURE

An important criterion for this controversy can be found from Faraone, Tsuang, and Tsuang (1999) who have used the term "genetically crisp" to define psychological measures that represent the homogenous effect of genes. A measure would be genetically crisp if the construct that it tapped into is affected primarily in a uniform way by a single set of genes. Faraone et al. (1999) argue that measures that have this characteristic are particularly important in applied situations because genetic counseling for disorders and molecular searches for candidate genes that lead to abnormal development depend on the uniform genetic composition of measures to be effective and reliable. Assessments based

on measures that are influenced by a single set of allelic variants will make more consistent diagnoses than judgments using a measure that is influenced by several sets of allelic variants which may or may not all be present in a given individual.

Although not discussed by Faraone et al. (1999), it may be similarly important for personality researchers to use genetically crisp measures, as this allows for more precise inference of a given psychological construct from a specific task. If an outcome is found to be linked to a broad personality factor that is an aggregate of several lower-order facets that have unique genetic variance, then it is difficult to determine what portion of the factor the outcome is associated with in actuality. Related to this point, increased statistical power is likely to accompany searches for allelic correlates of traits that are measured by instruments assessing uniform genetic effects. In other words, both candidate-gene and genome-wide association studies of personality are likely to be most successful when the personality outcome measured reflects a set of constructs that are influenced by a common set of genes, rather than a constellation of constructs, each of which is influenced by a unique set of genes. Research conducted at the phenotypic level is unable to provide information about the genetic crispness of different personality measurements. For example, even though Costa and McCrae (1995) found substantial variation that was specific to each facet, it is conceivable that all of the specific variance was due to environmental factors (and measurement error), and the higher order factor did capture all of the genetic variance of the lower-order traits. This would be evidence that aggregation to the FFM level does not lose significant genetic variance. However, if there is evidence for facet specific genetic variance that is not accounted for by the common factor, then this would lend support for narrow measurement.

There is a strong history of quantitative behavior genetic work within personality research. Behavior genetic analyses decompose the observable variance of a trait into

variance accounted for by additive genetic effects (A), shared environmental effects (C) that serve to make children from the same family similar to one another, and nonshared environmental effects (E) that are uncorrelated with family effects and make children different from one another. Analyzing personality factors and facets in this way can begin to uncover the etiology of traits which is a crucial task for constructing personality theory. The usefulness of the univariate behavior genetic approach has been demonstrated by removing at least one controversy from personality research. Some early conceptions of Agreeableness, Conscientiousness, and Openness to Experience defined these traits as learned tendencies to be contrasted with the biological traits of Extraversion and Neuroticism (Carver, Sutton, Scheier, 2000). Univariate behavioral genetic work has found this to be a false distinction, indicating fairly even contributions of both genes and environments to each of the five broad personality traits (Jang, McCrae, Angleitner, Riemann, & Livesley, 1998; McCrae & Costa, 1997; McCrae, Terracciano, & Members of the Personality Profiles of Cultures Project, 2005; Yamagata et al., 2006; Bergeman et al., 1993; Bouchard, 1997; Loehlin & Nichols, 1976; Loehlin, 1992; Riemann, Angleitner, & Strelau, 1997) as well as to the FFM facets (Jang, Livesley, & Vernon, 1996). This body of literature represents several age groups, several cultures, and a wide range of measures and recruitment strategies enhancing the generalizability and reliability of the findings. Furthermore, genetic influence has been found for the higher-order Stability and Plasticity factors (Jang et al., 2006) and for the general factor of personality (Veselka, Schermer, Petrides, & Vernon, 2009).

Such a univariate approach does little to advance the debate about the proper level of analysis, however, because univariate methods cannot make claims about the structure of genetic or environmental effects. Multivariate methods are necessary to distinguished between the commonality and specificity of genetic and environmental effects, and to test

whether common effects can be plausibly specified to operate indirectly on specific facets by way of more general traits. An early example of this approach comes from Heath and Martin (1990), who tested the assumption that phenotypic correlation and univariate estimates of heritability are evidence for coherent genetic and environmental effects within the Psychoticism trait (Eysenck & Eysenck, 1976). Based on data from 2,903 twin pairs, the researchers found that the phenotypic unity of the scale broke down into two separate genetic factors instead of one. In hindsight, this result is consistent with the FFM, as research has since indicated that Psychoticism represents a mixture of low Agreeableness and Conscientiousness (McCrae & Costa, 1985).

Univariate approaches to estimating genetic and environmental impacts on individual personality traits are now quite common, but there continues to be surprisingly little research that utilizes a multivariate approach. The few studies that have used a multivariate approach to examine the structure of the FFM have produced some conflicting results. McCrae, Jang, Livesley, Riemann, and Angleitner (2001) factor analyzed the genetic covariance structure of facets of the FFM in a sample of twins and uncovered five genetic factors. However, the genetic factors only roughly resembled the FFM. Yamagata et al. (2006) also point out that a methodological limitation (use of simple subtraction instead of the more precise structural equation modeling) of the study renders the findings equivocal: Did this limitation introduce noise into the data that created the lack of correspondence or did the genetic factors truly not represent the FFM? Support for the FFM comes from two separate groups of researchers (Yamagata et al., 2006; Pilia, et al., 2006) that have conducted more advanced exploratory factor analyses on the genetic and environmental correlation matrices of the FFM facets. They have found that the genetic and environmental factor loadings highly resemble the phenotypic loadings in that facets load highly on the related trait. While these approaches offer

strong evidence that the facets of the FFM are related at the phenotypic, genotypic, and environmental levels, the use of exploratory factor analysis is less able to provide accurate information about specific structural hypotheses such as trait hierarchical relationships which are best tested using confirmatory methods.

To our knowledge, only two research studies have utilized confirmatory, multivariate, behavior genetic methods to analyze the genetic and environmental structure of the FFM facets. Jang, Livesley, Angleitner, Riemann, and Vernon (2002) were the first to use confirmatory behavior genetic structural equation modeling methodologies to investigate the structure of the FFM. These researchers found that two, rather than one, genetic and nonshared environment latent factors were required to successfully account for the variance in the facets of each trait. Johnson and Krueger (2004) also found more complex genetic and environmental effects than expected by the FFM. Models that were fit with trait hierarchy as hypothesized by the FFM fit more poorly than more complex models for all traits except Extraversion and Neuroticism. These findings call into question whether the Big Five or any higher-order metatraits can be considered genetically crisp.

In addition to the importance of understanding the genetic and environmental structure of personality for all types of behavioral prediction, genome-wide association studies that investigate molecular behavior genetics are particularly likely to be advanced by multivariate quantitative behavior genetics. A few large scale and quite costly projects have been undertaken to search for genetic variants that might account for some variance in personality traits (Kim & Kim, 2011). It is common for these studies to find significant associations between specific genetic variants and broad personality traits, but for these associations to not replicate in separate samples (Terracciano et al., 2010; de Moor, et al., in press). These inconsistent results may stem from analyzing traits at too broad a level.

Terracciano and colleagues (2010) point out that imprecision of measurement renders it impossible to determine if an individual scored high on Neuroticism, for example, because they were high on Depression, but not Anxiety, or high on Anxiety, but not Depression. This imprecision of measurement is then passed down to imprecision in the search for specific genes that influence the trait significantly. Further, they argue that understanding the genetic structure of personality traits is an extremely valuable research goal as the DSM-V is moving towards a dimensional approach that conceptualizes psychiatric disorders as extremes along an otherwise normal-range distribution. Under such dimensional conceptualizations, determining what level of analysis should be emphasized in studies of normal-range personality may have implications for clinical diagnosis. If the most consistent and largest effects of genetic variants on personality outcomes occur at the levels of specific facets, then aggregation across facets may serve to dilute power in both molecular genetic research and applied practice. Alternatively, if genetic variants act at the trait level and environments contribute primarily to facet-level differentiation, then aggregation of scores across facets into larger traits may serve to maximize the genetic variation of interest by averaging out facet-specific environmental variation. In such a case, a high bandwidth approach may be most advantageous. Of course, it may be the case that regardless of the level of measurement, individual genes do not account for a substantial portion of any complex personality measure. It may be that genetic variation in personality results from a complex series of dynamic interactions between different biological, psychological, and environmental systems over the course of development (Turkheimer, 2000). Nevertheless, future research on such complex systems is likely to benefit from determination of the extent to which the processes occur at broad and specific levels of personality.

GOALS OF THE PRESENT STUDY

The present study seeks to extend and further explore the structure of the genetic and environmental effects on personality by fitting a series of structural equation models that specify varying degrees of structure for trait hierarchy. To fully understand the development of any characteristic, it is important to explicate the genetic and environmental effects that produce it. However, in light of the desire to create genetically crisp measures for both theoretical and applied goals, we place particular emphasis on the structure of the genetic effects on the facets. We will seek to provide a descriptive viewpoint of the relative importance of global and specific genetic effects for each domain of the FFM, as well as test differences between theoretically meaningful models. Evidence of the amount of genetic variance that is either shared among the facets within a domain or specific to individual facets can inform the discussion of whether important information about personality is contained or lost when aggregating to the domain level. Differences in the fit of psychometrically meaningful independent and common pathway models can shed light on whether the structure of the variance that is shared among facets within a domain can be represented in a highly coherent manner, such as a hierarchical factor, or is better represented by a looser interpretation (Neale & Cardon, 1992; McArdle & Goldsmith, 1990). Exploring the structure at the level of facets and domains can provide the necessary foundation for interpretations of the unity or coherence of even higher order metatraits or general factors of personality. Further replication of complex genetic and environmental effects on the facets with a novel, large sample of twins would shed light on the pragmatic use of facet level measurement.

METHOD

Participants

This study made use of data from a sample of adolescent twins who took part in the National Merit Twin Study (Loehlin & Nichols, 1976). The National Merit Twin Study was based on a sample of 596,241 students who took the National Merit Scholarship Qualifying Test (NMSQT) and answered an item asking if they have a twin. Based on this response, 1,507 potential same-sex pairs were identified and sent a questionnaire, of whom 850 twin-pairs ultimately participated (514 MZ and 336 DZ). Zygosity was determined based on a questionnaire developed by Nichols and Bilbro (1966). The final sample was composed of 354 (217 MZ and 137 DZ) male same-sex twin-pairs and 496 (297 MZ and 199 DZ) female same-sex twin-pairs. The sample used in the current study made use of the 807 twin-pairs for whom California Psychological Inventory (CPI) responses were available. The sample composition was 326 (202 MZ and 124 DZ) male same-sex twin-pairs and 481 (288 MZ and 193 DZ) female same-sex twin-pairs. Nearly all of the individuals who took the NMSQT were high school juniors. The majority of the individuals in the sample were therefore approximately 17 years of age. Note that while the students who take the NMSQT are higher achievers than a truly representative sample, the sample has been shown to be representative of the general population in terms of personality traits (Loehlin & Nichols, 1976). Further discussion of the sample and data collection procedures can be found in Loehlin and Nichols (1976).

Measures

Participants completed the CPI as part of a battery of tests (Gough, 1957). The CPI was constructed many years before the FFM was established. While the original CPI scoring system produces scales that appear related to Extraversion (e.g. Sociability and

Social Presence) and Conscientiousness (e.g. Responsibility or Self-Control), previous studies have found that the original CPI scales do not directly map onto the FFM fully (McCrae, Costa, & Piedmont, 1993). Fortunately, the CPI contains 480 items that the creators of the measure encourage to be recombined in different ways to construct new scales (Gough & Bradley, 1996). In this spirit Soto and John (2009) sought to extract the FFM from an item level analysis of the CPI. Using three independent validation samples, the researchers developed an algorithm by which to score the CPI on the FFM traits and several sub-facets. Each CPI item was organized into a cluster based on its primary correlation with two FFM traits. Then, the items were plotted into circumplex space based on the primary loadings. Groups of clustered items representing the different facets were identified within the circumplex space. The CPI-Big Five measure was found to be similar in each sample to several current measures of the FFM such as the Revised NEO Personality Inventory (Costa & McCrae, 1992a), the Big Five Inventory (John, Donahue, & Kentle, 1991), the Mini-Markers (Saucier, 1992), and the Adjective Check List (Gough & Heilbrun, 1983) as well as peer ratings on these measures. Across three samples, reliability estimates of the 16 facets averaged .72 and ranged from .56 to .85. Table 1 provides an overview of the 16 facets, and includes sample items. The current study utilized Soto and John's (2009) algorithm for scoring the CPI-Big Five.

Table 1. Facets and sample items of the CPI-Big Five

Scale	Paraphrased Sample Items
Extraversion	
Gregariousness	Enjoy gatherings where I can be with people. Like to have lots of friends and an active social life.
Social confidence vs. anxiety	Like talking in front of groups of people. Am nervous when meeting new people. (R)
Assertiveness	Have a ‘‘strong” personality. When in a group, usually do what others want. (R)
Agreeableness	
Trustfulness vs. cynicism	Am on my guard around very friendly people. (R) Believe that people pretend to care more than they actually do. (R)
Compassion vs. insensitivity	Would give money to right a wrong. Believe that people don’t need to worry about others. (R)
Humility vs. arrogance	Am likely to show off, when given the chance. (R) Sometimes act like I know more than I do. (R)
Conscientiousness	
Industriousness	Am a reliable worker. Do as little work as I can get by with. (R)
Orderliness	Like to have everything in its place. Find that planning takes most of the fun out of life. (R)
Self-discipline vs. distractibility	Find it hard to keep my mind on one thing. (R) Give up easily when I encounter problems. (R)
Neuroticism	
Anxiety	Worry about many things. Have very few fears. (R)
Depression	Am not as happy as others seem to be. Think that the future seems hopeless.
Rumination	Am often bothered by useless thoughts. Have bad habits that I can’t fight.
Irritability	Am bothered when unexpected events happen. Am sometimes grouchy without reason.
Openness	
Idealism vs. conformity	Have tried my hand at poetry. Believe that people should conform to the people around them. (R)
Intellectualism	Find the idea of research appealing. Don’t particularly enjoy learning new things. (R)
Adventurousness	Have had unusual experiences. Never do something for the thrill of it. (R)

Note: Table created after Table 1 of Soto and John (2009).

Analytical Approach

Our analytic approach rests on a comparison of two structurally meaningful models that differ in their allocation of common genetic and environmental effects on the facets of each Big Five domain. Figure 1 presents a graphical representation of these models (Neale & Caron, 1992; McArdle & Goldsmith, 1990). The first, depicted in Figure 1 A, is the independent pathways model. This model allows common genetic influences to directly affect each of the facets, rather than requiring them to influence the facets indirectly through a higher-order factor. This model allows genetic and environmental influences to have effects on multiple facets, but it does not presume these common effects to be manifestations of singular effects on a broader trait. The second model, depicted in Figure 1 B, is a common pathways model which places increased structure on the genetic and environmental effects. This model posits genetic and environmental effects on a latent variable, which are filtered down to the facets. Put differently, the common pathways model requires that the common genetic and environmental effects on a facet are proportional in terms of the facet's loading on the latent construct. Importantly, the primary distinction between these models is the amount of hierarchical structure that is posited for the common or global genetic and environmental effects. Both the independent and common pathways models specify unique or specific genetic and environmental effects on the facets that are not related to the effects of other facets. By global, we mean the variance that is shared amongst the facets within a domain. This is represented in Figure 1 as the variance components that are above the facets and are shared in common by each facet. By specific genetic and environmental variance, we mean the remaining variance within a facet that is not accounted for by the global variance and is represented in Figure 1 as the variance components below the facets that are unique to each facet.

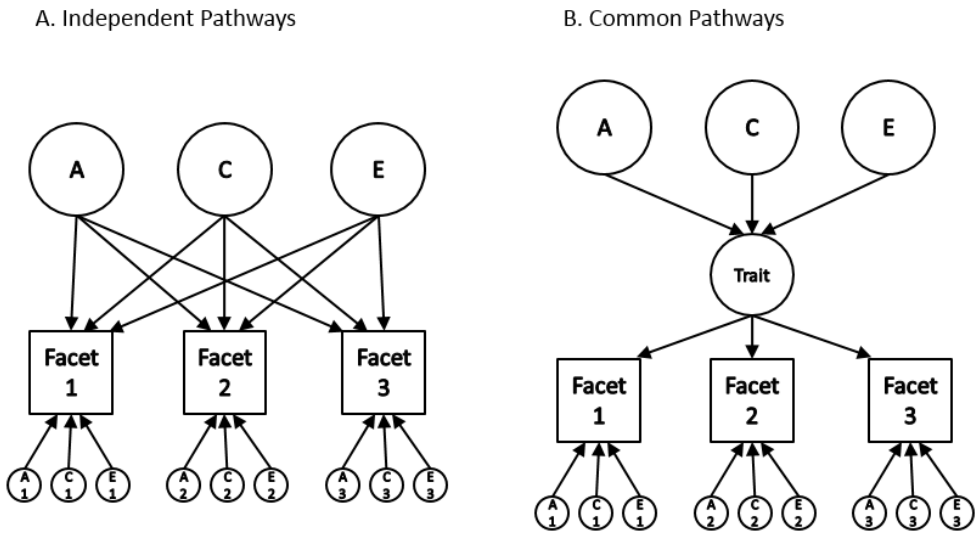

Figure 1. A. Diagram of independent pathways model. Global and specific ACE variables are allowed to directly affect the facets. **B.** Diagram of common pathways model. Global ACE variables are filtered through a hypothesized latent trait variable, and specific ACE variables are allowed to directly affect the facets.

We had two goals in mind for our approach to analyzing the data, a descriptive and a theoretical goal. The first goal was to simply describe the global genetic and environmental effects and the specific genetic and environmental effects on each trait construct. We report parameter estimates from the independent and common pathways models with all A, C, and E influences modeled. This allowed us to examine the magnitude of the variance components for each facet that are accounted for by broad factors that are shared with the other facets and the extent to which there are facet-specific, unique effects. Our intention is to better describe the different parameters instead of test their statistical significance, and we therefore report 95% confidence intervals for the estimates. Because the common pathways model is increasingly restrictive of the parameter estimates, the independent pathways model provides

estimated values that are closer to that implied by the data and may indicate the distinction between global and specific variance more accurately. However, comparing parameter estimates between the independent and common pathways models allows one to examine the influence that positing a hierarchical latent trait has on the genetic and environmental effects.

Second, our theoretical goal was to test the hypothesis that a latent phenotype can account for the genetic and environmental effects on facet scores without producing significant model misfit. The common pathways model is the more parsimonious model because it estimates fewer parameters. According to McArdle & Goldsmith (1990), the common pathways model is a nested form of the independent pathways model. We verified that this was the case using the procedure developed by Bentler and Satorra (2010). Our primary method of model comparison was therefore the χ^2 difference test. If two models did not fit significantly different from one another according to the χ^2 difference test, the more parsimonious common pathways model was preferred. Also provided are Akaike's Information Criterion (AIC) and root mean square error of approximation (RMSEA). AIC reflects a balance between parsimony and fit and can be used to compare both nested and non-nested models. Lower AIC values indicate better fit. RMSEA is an index of absolute model fit and is not used for direct comparison. RMSEA values below .05 indicate good model fit between the expected and observed covariance matrix (Steiger, 1989).

All models were fit using full-information maximum-likelihood estimation in M*plus* statistical software (Muthén & Muthén, 1998-2010).

RESULTS

Scale Description

Our first step was to determine if the CPI-Big Five scoring algorithm captured the FFM in the new sample using exploratory factor analysis of the produced facet scores. Separate analyses were conducted for the first and second member of a twin pair (twin1 and twin2). We used orthogonal target rotation to maximize the correspondence of this structure to the factor loadings reported by Soto and John (2009). The first five eigenvalues were 4.08/4.23, 2.55/.257, 1.65/1.71, 1.51/1.40, and 1.06/1.02 followed by .78/.70, .63/.65, .55/.53, .50/.49, and .48/.47 (for twin1 and twin2 respectively). From the eigenvalues, we determined that the expected five factor structure would be retained. Table 2 presents the factor loadings for both twins, as well as factor and variable congruence coefficients based on the congruence with the loadings reported by Soto and John (2009). Factor congruence coefficients ranged from .93 to .97, and variable congruence coefficients ranged from .86 to 1.00. Of 32 possible primary loadings, 29 facet primary loadings were for their intended trait. The one exception for both twin1 and twin2 was the facet Adventurousness which loaded most strongly (and negatively) on Agreeableness, rather than on Openness, as found by Soto and John (2009). This may indicate that the facets that indicate Openness in this sample, as takers of the NMSQT, may not be representative of the general population or more specifically, the student and community samples utilized by Soto and John (2009). Another difference was found for Compassion vs. Insensitivity which loaded slightly stronger on Openness than Agreeableness for twin2. Together, this misfit at the phenotypic level may indicate that the results for these traits should be interpreted with some caution. However, the overall structure is largely similar as indicated by the factor and variable congruence coefficients, and the factor structure indicated by Soto and John (2009) was retained

Table 2. Target factor loadings of the CPI-Big Five in current and previous study

Scale	E	A	C	N	O	VCC
Gregariousness	**.72/.82**	-.17/-.22	.01/-.04	-.14/-.10	.01/-.03	.99/.97
Social confidence vs. anxiety	**.60/.61**	-.28/-.29	.24/.26	-.31/-.30	.24/.27	.92/.98
Assertiveness	**.59/.50**	-.34/-.39	.22/.33	-.23/-.23	.31/.33	.98/.97
Trustfulness vs. cynicism	.21/.16	**.71/.66**	.13/.18	-.38/-.39	.12/.07	.99/.98
Compassion vs. insensitivity	.16/.21	**.38**/.36	.03/-.04	-.03/.14	.35/**.40**	.96/.97
Humility vs. arrogance	-.07/-.14	**.57/.54**	.23/.20	-.10/-.19	-.14/-.08	.86/.86
Industriousness	.07/.05	.20/.18	**.72/.69**	-.08/-.11	.01/-.07	.99/.98
Orderliness	.02/.14	.02/.18	**.70/.63**	-.08/-.11	-.23/-.23	.96/.93
Self-discipline vs. distractibility	.11/.05	.14/.12	**.53/.46**	-.36/-.39	.06/.10	.97/.97
Anxiety	-.01/-.10	.03/-.04	.05/.05	**.81/.73**	-.12/-.05	1.00/.98
Depression	-.23/-.17	-.15/-.15	-.18/-.22	**.70/.73**	-.05/-.05	.97/.97
Rumination	-.10/-.12	-.28/-.33	-.27/-.22	**.52/.58**	.09/.11	.99/.99
Irritability	-.18/-.13	-.30/-.31	-.06/-.14	**.54/.58**	-.19/-.18	.95/.92
Idealism vs. conformity	.02/.00	.17/.20	-.02/.01	.12/.12	**.66/.68**	.98/.98
Intellectualism	-.04/.00	-.10/-.07	.38/.31	-.17/-.24	**.64/.63**	.98/.99
Adventurousness	.11/.15	**-.49/-.47**	-.31/-.39	.05/.06	.39/.32	.94/.91
Factor Congruence Coefficient	.97/.96	.97/.97	.95/.93	.97/.97	.97/.96	

Note: Values are for twin1 followed by twin2. VCC is the variable congruence coefficient. Congruence coefficients estimate the agreement between the loadings for each member of the twin pair and the loadings as reported by Soto and John (2009). Primary facet loading values are printed in bold.

Table 3 presents the univariate estimates and confidence intervals of the effects of genes, shared environment, and nonshared environment on the facets and on composites representing the five broader domains. In line with previous research (e.g., Loehlin, 1992), estimates of heritability were typically moderate to large and ranged from .13 to .63 with a mean of .40 and median of .37 for the facets. Also in line with previous research (Bouchard, 1997), the facets showed null-to-small effects of the shared environment (ranging from 0 to.23, mean = .08, median = .05). The remaining facet variance was explained by nonshared environment with estimates that were moderate to large in magnitude (ranging from .38 to .65, mean = .53, median = .56).

Table 3. Univariate estimates of the effect of genes and the environment on CPI-Big Five traits and facets

Scale	h^2 [95% CI]	c^2 [95% CI]	e^2 [95% CI]
Extraversion	.48 [.29 - .71]	.12 [.01 - .36]	.41 [.36 - .46]
Gregariousness	.55 [.36 - .78]	.02 [0 - .70]	.43 [.38 - .49]
Social confidence vs. anxiety	.59 [.48 - .60]	.00 [0 - .66]	.46 [.40 - .52]
Assertiveness	.58 [.53 - .63]	.00 [0 - .78]	.42 [.37 - .48]
Agreeableness	.40 [.20 - .66]	.12 [0 - .41]	.48 [.42 - .53]
Trustfulness vs. cynicism	.18 [.02 - .51]	.20 [.04 - .46]	.62 [.55 - .70]
Compassion vs. insensitivity	.28 [.08 - .59]	.09 [0 - .44]	.63 [.56 - .71]
Humility vs. arrogance	.46 [.40 - .53]	.00 [0 - .95]	.54 [.48 - .61]
Conscientiousness	.46 [.40 - .53]	.00 [0 – 1]	.53 [.46 - .61]
Industriousness	.43 [.25 - .67]	.10 [0 - .38]	.47 [.41 - .53]
Orderliness	.31 [.11 - .61]	.05 [0 - .52]	.65 [.57 - .73]
Self-discipline vs. distractibility	.37 [.30 - .45]	.00 [0 - .13]	.63 [.56 - .70]
Neuroticism	.40 [.20 - .66]	.11 [0 - .41]	.49 [.42 - .56]
Anxiety	.47 [.28 - .71]	.10 [0 - .39]	.43 [.37 - .49]
Depression	.37 [.17 - .66]	.05 [0 - .49]	.58 [.51 - .65]
Rumination	.13 [0 - .49]	.23 [.07 - .49]	.64 [.57 - .72]
Irritability	.33 [.13 - .62]	.03 [0 - .57]	.64 [.56 - .72]
Openness	.53 [.36 - .74]	.13 [.01 - .37]	.34 [.29 - .38]
Idealism vs. conformity	.41 [.24 - .63]	.20 [.07 - .42]	.38 [.33 - .43]
Intellectualism	.32 [.15 - .57]	.21 [.07 - .44]	.46 [.40 - .53]
Adventurousness	.40 [.19 - .67]	.05 [0 - .48]	.57 [.49 - .63]

Note: 95% CI is the 95% confidence interval. Confidence intervals were constrained to the plausible limits of the proportion of variance that could be accounted for (i.e., ranging from 0 to 1). h^2 is the proportion of variance attributable to additive genetic effects, c^2 is the proportion of variance attributable to shared environmental effects, and e^2 is the proportion of variance attributable to nonshared environmental effects.

Multivariate Parameter Estimation

Table 4 presents the parameter estimates and 95% confidence intervals from the ACE independent pathways model for each facet. A number of features of are note. First, there is strong evidence for common genetic effects across facets within a domain as evidenced by the high global genetic effect parameter estimates ranging from .18 to .71 (mean = .48, median = .48). Second, there is also strong evidence for specific genetic effects that operate on the individual facets and are not shared across facets. Estimates of specific genetic effects ranged from 0 to .57 (mean = .29, median = .31). This observation indicates that aggregating across facets within a trait would necessarily result in loss of specific genetic variance for many facets. Third, it can be seen that environmental effects are also fairly evenly distinguished into global and specific influences with a few exceptions. Global shared environmental estimates ranged from .02 to .52 (mean = .21, median = .19), and specific estimates ranged from 0 to .39 (mean = .12, median = 0). Finally, global nonshared environmental parameter estimates ranged from .03 to .75 (mean = .41, median = .41), and specific nonshared environmental parameter estimates ranged from 0 to .79 (mean .54, median .57).

Table 4. Standardized Parameter estimates and variance components from the independent pathways model of each facet

Scale	Parameter Point Estimates [95% Confidence Interval]					
	A^C	C^C	E^C	A^S	C^S	E^S
Gregariousness	.47 [.31 - .63]	.14 [0 - .80]	.41 [.35 - .48]	.57 [.48 - .66]	.00 [0 - .75]	.52 [.47 - .56]
Social confidence vs. anxiety	.71 [.59 - .84]	.09 [0 - .49]	.49 [.43 - .56]	.14 [0 - .49]	.00 [0 - .50]	.47 [.41 - .52]
Assertiveness	.64 [.51 - .76]	.08 [0 - .50]	.43 [.37 - .49]	.41 [.30 - .51]	.00 [0 - .54]	.49 [.44 - .54]
Trustfulness vs. cynicism	.43 [.20 - .65]	.45 [.25 - .64]	.75 [.18 - 1]	.00 [0 - .54]	.00 [0 - .45]	.24 [0 – 1]
Compassion vs. insensitivity	.54 [.33 - .75]	.04 [0 - .34]	.10 [0 - .20]	.00 [0 – 1]	.28 [0 - .66]	.79 [.74 - .83]
Humility vs. arrogance	.18 [0 - .38]	.33 [.20 - .46]	.27 [.05 - .49]	.55 [.49 - .61]	.00 [0 - .34]	.69 [.60 - .78]
Industriousness	.57 [.39 - .75]	.15 [0 - .53]	.41 [.33 - .49]	.27 [0 - .67]	.32 [.09 - .55]	.55 [.50 - .61]
Orderliness	.42 [.24 - .60]	.30 [0 - .60]	.52 [.42 - .61]	.29 [0 - .64]	.00 [0 – 1]	.62 [.55 - .69]
Self-discipline vs. distractibility	.46 [.26 - .65]	.06 [0 - .33]	.40 [.32 - .49]	.40 [.16 - .64]	.00 [0 - .37]	.68 [.63 - .74]
Anxiety	.65 [.57 - .74]	.02 [0 - .33]	.40 [.34 - .45]	.33 [.04 - .61]	.22 [0 - .55]	.51 [.47 - .56]
Depression	.47 [.33 - .61]	.22 [0 -.45]	.60 [.54 - .66]	.34 [.14 - .55]	.15 [0 - .54]	.48 [.42 - .54]
Rumination	.27 [0 - .55]	.52 [.39 - .66]	.50 [.43 - .57]	.00 [0 - .85]	.00 [0 – 1]	.64 [.59 - .68]
Irritability	.48 [.36 - .59]	.14 [0 - .37]	.39 [.32 - .47]	.25 [0 - .62]	.23 [0 - .57]	.70 [.65 - .74]
Idealism vs. conformity	.48 [.23 - .72]	.24 [0 - .57]	.10 [.04 - .15]	.43 [.12 - .73]	.39 [.09 - .68]	.61 [.57 - .65]
Intellectualism	.52 [.26 - .78]	.30 [0 - .70]	.68 [.64 - .73]	.24 [0 - .85]	.35 [0 - .77]	.00 [0 – 1]
Adventurousness	.41 [.20 - .62]	.28 [.03 - .53]	.03 [0 - .10]	.44 [.14 - 74]	.00 [0 - .99]	.75 [.70 - .79]

Note: Confidence intervals were constrained to the plausible limits of standardized parameter estimates for a behavior genetic model (i.e., ranging from 0 to 1). A^C represents the common effect of genes; C^C represents the common effect of shared environment; E^C represents the common effect of nonshared environment; A^S represents the effect of genes specific to the facet; C^S represents the effect of shared environment specific to the facet; E^S represents the effect of nonshared environment specific to the facet.

The proportion of total genetic, shared environmental, and nonshared environmental effects that are accounted for by global or common effects is of particular interest for the current study. Table 5 presents the proportion of each total effect, the sum

of the common and unique variance attributable to A, C, or E, which is due to common effects as implied by the parameter estimates from the independent pathway model. This provides an estimate of the relative importance of global or specific effects for each variance component. For example, relative to the total effect of genetics on Depression, 65% of the effect is shared with other facets of Neuroticism. Alternatively, 100% of the effect of genetics is accounted for by global effects shared between the facets for Trustfulness vs. Cynicism, Compassion vs. Insensitivity, and Rumination. This indicates that aggregating these facets to the trait level captures all of the genetic information found at the facet level. However, these facets are the exception. For Extraversion facets, the average percent of the genetic effect that is accounted for globally is 69%. Very little specific information is lost for Social Confidence vs. Anxiety (96% global), but the majority of the information is lost for Gregariousness (40% global). Two facets of Agreeableness show complete genetic overlap at the common level, but Humility vs. Arrogance showed almost no genetic relation to the other facets (9% global). There is strong evidence of common genetic influences for Conscientiousness and Neuroticism facets with all facets sharing more than half of the genetic variance globally and averaging 69% and 81% global genetic variance, respectively. The Intellectualism facet strongly defined the genetic component of Openness with 83% of its genetic variance accounted for globally, but only 46% and 55% of the genetic variance in Idealism vs. Conformity and Adventurousness is shared among the facets. Similar interpretations can be made for the shared and nonshared environment. While there is some evidence that genetic information is represented completely in an aggregate trait, it is typically the case that substantial information remains unique to each facet.

Table 5. Proportion of genetic, shared environmental, and nonshared environmental variance account for by global effects

Scale	A	C	E
Extraversion			
Gregariousness	.40	1.00	.39
Social confidence vs. anxiety	.96	1.00	.53
Assertiveness	.71	1.00	.43
Agreeableness			
Trustfulness vs. cynicism	1.00	1.00	.91
Compassion vs. insensitivity	1.00	.02	.02
Humility vs. arrogance	.09	1.00	.13
Conscientiousness			
Industriousness	.81	.19	.38
Orderliness	.68	1.00	.41
Self-discipline vs. distractibility	.57	1.00	.26
Neuroticism			
Anxiety	.80	.01	.37
Depression	.65	.68	.61
Rumination	1.00	1.00	.38
Irritability	.78	.28	.24
Openness			
Idealism vs. conformity	.55	.29	.02
Intellectualism	.83	.42	1.00
Adventurousness	.46	1.00	.00

Note: Values calculated from the estimates of the independent pathway model by dividing the variance accounted for globally in A, C, or E by the total effect of A, C, or E.

Table 6 reports the parameter estimates and 95% confidence intervals for the more constrained common pathways model. These estimates indicate strong common

genetic effects for each of the broad traits with point estimates for the common influence of A ranging from .53 to .89 (mean = .74, median = .76), and strong common nonshared environmental effects for each of the traits, with point estimates ranged from .39 to .68 (mean = .60, median = .65). There is less evidence for the effects of the shared environment, but there is much variation between traits with estimates ranging from .00 to .50 (mean = .20, median = .23). Also of importance for these analyses, much specific genetic (mean = .32, median .39) and nonshared environmental (mean = .60, median = .60) variance remains at the facet level with some evidence of shared environmental (mean = .19, median = .24) variance. These results are generally consistent with those of the independent pathways models, in indicating that a great deal of genetic variation is facet-specific. Of course, because it imposes a highly constrained structure on the patterns of common genetic and environmental variation in the factors, its parameters may be less trustworthy than those from the independent pathways model. To evaluate whether the more constrained common pathways model represents the data as well as the less constrained independent pathways model, we turned to model comparison tests.

Table 6. Standardized Parameter estimates from the common pathways model of each facet

	Parameter Point Estimates [95% Confidence Interval]			
Scale	Factor Loading	A	C	E
		.82	.00	.58
Common		[.78 - .86]	[0 - 1]	[.53 - .63]
	.64	.57	.00	.52
Specific Gregariousness	[.60 - .68]	[.52 - .63]	[0 - 1]	[.48 - .56]
Specific Social confidence	.87	.16	.00	.46
vs. anxiety	[.84 - .90]	[.01 - .31]	[0 - .46]	[.41 - .51]
	.77	.46	.00	.48
Specific Assertiveness	[.74 - .80]	[.40 - .52]	[0 - .48]	[.44 - .52]
		.53	.50	.68
Common		[.24 - .82]	[.24 - .76]	[.60 - .76]
Specific Trustfulness vs.	.86	.00	.00	.51
cynicism	[.80 - .92]	[0 - .21]	[0 - .19]	[.41 - .61]
Specific Compassion vs.	.32	.41	.34	.78
insensitivity	[.26 - .38]	[.12 - .70]	[.03 - .65]	[.73 - .83]
Specific Humility vs.	.49	.57	.00	.66
arrogance	[.43 - .55]	[.52 - .62]	[0 - .36]	[.61 - .71]
		.76	.00	.65
Common		[.70 - .82]	[0 - .65]	[.58 - .72]
	.73	.32	.31	.51
Specific Industriousness	[.68 - .78]	[.06 - .58]	[.08 - .54]	[.45 - .56]
	.68	.23	.24	.66
Specific Orderliness	[.63 - .73]	[0 - .66]	[0 - .58]	[.61 - .71]
Specific Self-discipline vs.	.57	.45	.00	.69
distractibility	[.52 - .62]	[.38 - .52]	[0 - .42]	[.64 - .74]
		.70	.23	.68
Common		[.49 -.91]	[0 - .78]	[.62 - .74]
	.67	.50	.24	.49
Specific Anxiety	[.63 - .71]	[.35 - .65]	[0 - .51]	[.44 - .54]
	.82	.26	.10	.51
Specific Depression	[.79 - .85]	[0 - .53]	[0 - .71]	[.46 - .56]
	.64	.00	.41	.65
Specific Rumination	[.60 - .68]	[0 - .77]	[.35 - .47]	[.61 - .69]
	.63	.28	.24	.69
Specific Irritability	[.59 - .67]	[0 - .62]	[0 - .58]	[.64 - .74]
		.89	.25	.39
Common		[.69 - 1]	[0 - .93]	[.28 - .50]
Specific Idealism vs.	.58	.41	.41	.57
conformity	[.48 - .68]	[.16 - .66]	[.20 - .62]	[.52 - .63]
	.66	.01	.41	.63
Specific Intellectualism	[.54 - .78]	[0 - 1]	[.27 - .55]	[.57 - .69]
	.23	.55	.31	.74
Specific Adventurousness	[.16 - .30]	[.34 - .76]	[0 - .64]	[.69 - .79]

Note: Confidence intervals were constrained to the plausible limits of standardized parameter estimates for a behavior genetic model (i.e., ranging from 0 to 1).

Model Comparison

The independent pathways model represents a relatively unrestricted account of the patterns of global and specific genetic and environmental influences on the facets within each of the trait clusters. It is possible that a more parsimonious common pathways model (in which higher order factors are explicitly presumed) or a model that does not estimate all of the parameters can fit the data equally well. Following the precedent of Johnson and Kreuger (2004) and Jang and colleagues (2002), we compared models that did or did not posit the effect of the shared environment on the facets. The purpose of this is twofold. If models without the shared environmental effect specified do not fit significantly worsen, this more parsimonious model can be taken as the best representation of the data. Further, in some instances the MZ correlation was slightly more than double the DZ correlation which may imply the effect of dominant genes. Demonstrating that model fit is not significantly reduced by removing the effect of the shared environment allows for the effects of dominance to be practically modeled in place of the shared environment. Table 7 presents the multivariate model fit comparison statistics for all models under investigation. RMSEA values were below .05 for all models examined indicating that they all fit the data relatively well. Removing the effect of shared environment produced a significant increase in χ^2 for Neuroticism and Openness and did not for Extraversion, Agreeableness, and Conscientiousness. No models including the effect of dominant genes, however, fit significantly better than the AE models. Thus, for the purposes of comparing the common and independent pathways models, ACE Neuroticism and Openness models were evaluated, and AE Extraversion, Agreeableness, and Conscientiousness models were evaluated.

Table 7. Comparisons of ACE and AE models.

Factor and Model	χ^2	*df*	*p*	RMSEA	90% UL	χ^2_D	df_D	p_D
Extraversion								
ACE model	37.43	33	.27	.018	.042			
AE model	**37.48**	**39**	**.54**	**.000**	**.033**	**0.05**	**6**	**.99**
ADE model	37.06	33	.29	.017	.042	-0.42	6	.99
Agreeableness								
ACE model	44.32	33	.09	.029	.050			
AE model	**54.58**	**39**	**.05**	**.031**	**.050**	**10.26**	**6**	**.11**
ADE model	48.05	33	.04	.034	.053	-6.53	6	.37
Conscientiousness								
ACE model	30.52	33	.59	.000	.033			
AE model	**33.94**	**39**	**.70**	**.000**	**.028**	**3.42**	**6**	**.75**
ADE model	31.18	33	.56	.000	.034	-2.76	6	.84
Neuroticism								
ACE model	**84.00**	**60**	**.02**	**.031**	**.046**			
AE model	112.76	68	.00	.040	.053	28.76	8	.00
Openness								
ACE model	**26.14**	**33**	**.80**	**.000**	**.025**			
AE model	40.01	39	.42	.008	.036	13.87	6	.03

Note: All comparisons based on the independent pathways model and reflect change in χ^2 from the model immediately preceding it in the table. Positive χ^2 difference values indicate that model misfit occurs in the comparison, and negative χ^2 difference values indicate that model misfit has been removed by the model. Significant p_D values indicate that the fit of the model is significantly different. Preferred models are in bold.

Chi-squared difference tests were used to determine if the nested common pathways and independent pathways models fit significantly different (McArdle & Goldsmith, 1990). Table 8 lists the fit statistics for these comparisons. The models did not fit differently for Extraversion ($\chi^2_D = 2.70$, $df_D = 2$, $p = .26$) or Openness ($\chi^2_D = 5.02$,

$df_D = 4$, $p = .29$). Introducing the latent variable for Agreeableness produced a significant worsening of fit ($\chi^2_D = 8.87$, $df_D = 2$, $p = .01$). A similar worsening of fit was found for Conscientiousness ($\chi^2_D = 6.12$, $df_D = 2$, $p = .05$) and Neuroticism ($\chi^2_D = 35.52$, $df_D = 4$, $p < .001$). To summarize, an AE common pathways model was retained for Extraversion, an AE independent pathways model was retained for Agreeableness and Conscientiousness, an ACE independent pathways model was retained for Neuroticism, and an ACE common pathways model was retained for Openness.

Table 8. Comparisons of Independent Pathways and Common Pathways Models.

Factor and Model	χ^2	*df*	*p*	AIC	RMSEA	90% UL	χ^2_D	df_D	p_D
Extraversion									
IP AE model	37.48	39	.54	33614.24	.000	.033			
CP AE model	**40.18**	**41**	**.51**	**33612.93**	**.000**	**.033**	**2.70**	**2**	**.26**
Agreeableness									
IP AE model	**54.58**	**39**	**.05**	**34797.18**	**.031**	**.050**			
CP AE model	63.45	41	.01	34802.05	.037	.054	8.87	2	.01
Conscientiousness									
IP AE model	**33.94**	**39**	**.70**	**35091.11**	**.000**	**.028**			
CP AE model	40.05	41	.51	35093.23	.000	.033	6.12	2	.05
Neuroticism									
IP ACE model	**84.00**	**60**	**.02**	**45922.58**	**.031**	**.046**			
CP ACE model	119.52	66	.00	45946.10	.045	.058	35.52	6	.00
Openness									
IP ACE model	26.14	33	.80	34606.42	.000	.025			
CP ACE model	**31.16**	**37**	**.74**	**34603.44**	**.000**	**.026**	**5.02**	**4**	**.29**

Note: IP = Independent Pathways. CP = Common Pathways. Significant p_D values indicate that the fit of the model is significantly worse.

Model Misspecification Analysis

While the χ^2 difference tests reported above were capable of determining whether significant model misfit occurs when constraining the independent pathways model to the common pathways model, the tests do not offer information about where the misfit is occurring. There are a number of different modeling strategies available to evaluate and determine more desirable structural relations. One possible strategy would begin with the model that produced a significant increase in model misfit and relax some of the constraints or add parameters until acceptable fit to the data is obtained. Such a strategy would be particularly useful when these additions or modifications would have a priori theoretical or structural meaning. An alternative approach would be to comprehensively explore potential sources of misfit in a model without endorsing specific modifications to the model. Such a strategy is more concerned with describing sources of model misspecification and the plausible tenability of the restrictive model rather than an explicit attempt to elaborate on the model.

The current analysis took the latter of the above two approaches. We are not aware of any theoretical speculation about the facets utilized in this study in terms of special genetic or environmental effects between common variance components and facets, nor did we have any a priori expectations about this issue. Further, it is unclear what the theoretical significance of finding that, for example, the Industriousness facet of Conscientiousness requires an additional pathway from a global variance component to account for the component's influence above and beyond that mediated by the latent trait. A second, more pragmatic, concern for conducting a theoretically driven modeling strategy is the limited number of facets (3-4) and variance components (sometimes only two) that were used for model comparison limiting the number of possible modifications

that can be made. Together, these concerns are substantial enough that any analysis based solely on improving the misfitting models used for comparison may be misleading. Therefore, we sought to descriptively, rather than theoretically, explore potential sources of model misfit that are found in the common pathways model. This was accomplished by freeing the common pathways-induced proportionality constraint (see Appendix) from each A, C, and E component individually for each facet and noting the change in the χ^2 value. This is a one degree of freedom difference test and indicates the overall amount of improvement in fit the model obtains by allowing a single variance component to no longer be constrained to be proportional with the other variance components of a specific facet.

Figure 2 displays a bar graph of the improvement in model fit as measured by change in χ^2 by freeing one parameter. Of the 48 possibly significant improvements in model fit, only five reach the critical value for a one degree of freedom χ^2 test. No individual significant improvements were found for Extraversion, Openness, and Neuroticism. This was to be expected for Extraversion and Openness as the common pathways model did not fit significantly worse than the independent pathways model for these traits. The finding is less clear for the Neuroticism domain which produced a substantial increase in model misfit by imposing the common pathways model. The general trend is for small improvements in fit for each variance component for each facet. This indicates that rather than one variance component or a single facet being responsible for model misspecification, the decrements of model fit are due to minor, distributed effects. Models for Agreeableness and Conscientiousness that freed a single parameter displayed significantly improved fit, particularly for genetic and nonshared environmental effects. For Agreeableness, the majority of improvement was found for the facets of Trustfulness and Compassion, and for Conscientiousness, the majority of

improvement was found for the facet of Industriousness. These specific sources of misfit should be considered tentative possibilities for modifying the common pathways model until they are replicated in future work. Additionally, it is important to keep in mind that the improvements in fit indicated by individual comparisons are not additive improvements and may be overlapping. Due to these limitations and the descriptive nature of this analysis, we choose not to interpret these results further.

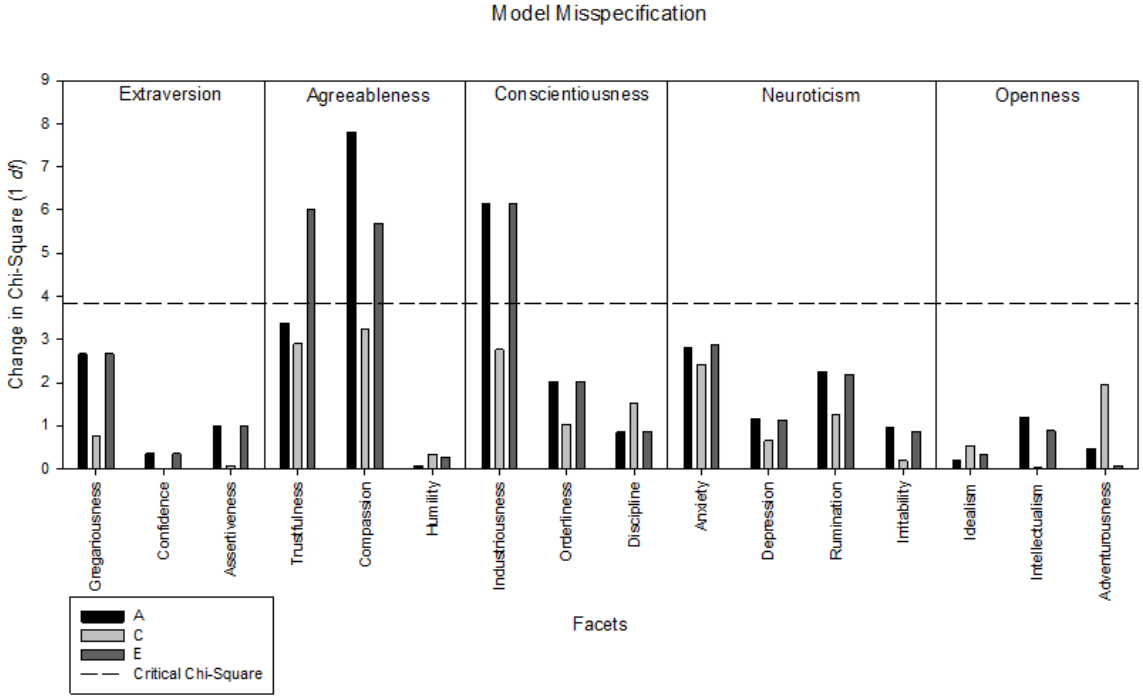

Figure 2. Bar graph of the influence of each facet's variance component parameters on improving model fit by sequentially freeing the proportionality constraint of the common pathways model.

DISCUSSION

How to interpret the structural organization of personality has long been a contentious argument in psychology. Behavior genetic research of personality structure takes up the challenge laid out by Eysenck (1992b) "to anchor our dimensions of personality in something more concrete than the morass of factor analysis" with an

understanding that "biology supplies us with the necessary tools" to build consistent theories (p. 672). By applying multivariate behavioral genetic models to data from MZ and DZ twins reared together, we found that there is significant fluctuation in the structure of genetic and environmental effects on personality traits. Of particular importance, we found evidence of statistically significant specific genetic influences on facets not accounted for by a common hierarchical trait as modeled in the common pathways model for ten of 16 facets. This implies that aggregating facet scores to produce domain scores overshadows the nuances the genetic effects and renders the domain scores not genetically crisp. The empirical story that the parameter estimates present is one of important and complex variation at the level of the facets that underlie broad domains.

Both the common and independent pathways models indicated common genetic influences on the facets within each domain, but comparison of differences in model fit indicated these common influences were not well accounted for by the operation of genetic and environmental influences on higher-order traits representing Neuroticism, Conscientiousness, and Agreeableness. The common pathways model did not fit significantly worse for Extraversion and Openness, which is consistent with a conceptualization of genetic and environmental effects being filtered through latent traits. Our results of a complex genetic and environmental structure of personality are generally consistent with the previous studies that have examined the factor structure of personality traits using biometric methods. Jang et al. (2002) found multiple genetic factors were required to explain the variance in the facets. This is similar to our finding that a simple latent trait perspective does not capture all of the common variance in personality constructs. Johnson and Krueger (2004) used similar common and independent pathways modeling techniques to examine their data. They found that the common pathways model

fit for Extraversion and Neuroticism. We found that the common pathways model fit for Extraversion and Openness. While this provides replication of the unitary nature of Extraversion, it casts doubt onto the replicability or structural cohesiveness for Neuroticism and Openness. Future research can help determine which results may be simply sample specific and which will be generalizable. For the remaining traits, the independent pathways model fit best. Even allowing for these inconsistencies, it is clear from the current and past results that the content and structure of personality facets goes beyond what is found at the higher-order factor level.

The utility of the FFM is in its capacity to comprehensively (or near to it) describe patterns of variation in the behavioral tendencies of individuals in five simple factors. The common practice of analyzing data on composite scores of each of the five domains may have caused interesting and important associations between lower-order facets and outcomes to be masked. Our results indicate that, to advance personality theory, nuanced viewpoints of differential effects of facets should be further explored. For instance, finding that at the domain level certain outcomes are not related to personality does not imply that there will be no associations at the facet level. Additionally, relationships that are found at the domain level may move closer to uncovering mechanisms for the relationship when examined at the facet level. One successful example of this can be found from Luciano et al. (2006) who explored the relationship between Conscientiousness at the facet level and measures of IQ and academic achievement. Only certain facets, notably Competence, Dutifulness, and Deliberation, were found to be significantly associated with intelligence or achievement. This finding at the phenotypic level led the researchers to explore the genetic relationship of these variables and find strong evidence of a common genetic component for the personality and ability measures.

This emphasizes the pragmatic utility of facet level measurement for uncovering unique genetic relationships between personality and other outcomes.

Based on the complex and somewhat disperse structure of the FFM should researchers focus away from higher order traits and towards specific facets instead? We would argue that all levels of analysis have important uses within personality psychology (see Costa & McCrae, 1995 and Paunonen & Ashton, 2001b for discussion of when more or less detail is desirable). Costa and McCrae (1992b) claimed that "there are neurobiological structures that underlie such heritable personality traits" (p. 659). While we do not believe that multivariate analyses, be they phenotypic or behavioral genetic, can be used to directly evaluate this speculation, we do note that the biological substrate of personality need not be highly centralized in order for the shared functions of biological structures to be adequately described by a broad trait. For example, genome-wide association studies have indicated that complex and distributed interactions between many biological structures are likely crucial to larger constructs (Terracciano et al., 2010).

Costa and McCrae (1992b) have argued for measurement at the level of the Big Five Factors from a pragmatic perspective. They cite the extensive literature on the real-world correlations between important outcomes and personality traits as evidence for their utility (e.g. Costa, McCrae, & Holland, 1984; Barrick & Mount, 1991; McCrae & Costa, 1991; Costa, McCrae, & Dembroski, 1989; Digman & Takemoto-Chock, 1981; Miller, 1991). The issue of pragmatic utility is distinguishable from that of biological or psychology structure. As Cronbach and Meehl (1955) wrote:

> Factors may or may not be weighted with surplus meaning. Certainly when they are regarded as 'real dimensions' a great deal of surplus meaning is implied, and the interpreter must shoulder a substantial burden of proof. The alternative view is

> to regard factors as defining a working reference frame, located in a convenient manner in the "space" defined by all behaviors of a given type. Which set of factors from a given matrix is "most useful" will depend partly on predilections, but in essence the best construct is the one around which we can build the greatest number of inferences, in the most direct fashion. (p. 287-288)

Indeed, our results suggest that while the Big Five personality factors may function as highly useful working reference frames at a broad level, additional important variation occurs at the facet level. One important confirmation of this argument is empirical findings of differential within-factor facet effects on outcomes (Paunonen, 1998). The widespread existence of these effects and inability of even the Big Five to act as latent traits that filter global genetic and environmental effects onto lower level facets raises serious concerns about the long-term benefits of studying higher-order factors, such as Stability and Plasticity or the general factor of personality. While these factors may be able to produce some replicable and consistent findings, their inability to account for the genetic variation among the facet-level traits renders fine grained analysis more productive for future research. It is important to note that this study did not empirically evaluate the plausibility of higher order factors accounting for the genetic variation in the facets, but we would argue that it can be inferred from the current results that the Big Five factors could not account for all of the genetic variation in the facets. Thus, metatraits or the general factor (which typically use the Big Five as a measurement portion of a model) would similarly be unable to account for the genetic variance in the facets. Further, that the common pathways model fit significantly worse for three traits is evidence that the higher order factors may not be as unitary as phenotypic results may indicate. Future research should investigate the extent to which the etiological structure found in the current study is altered by incorporating higher-order structure. Rather than a

broad general factor influencing a behavioral outcome, we would argue that, in terms of behavioral mechanisms and predictions as well as molecular searches, a more frequent consideration for facet-level measurement would enhance the pragmatic utility of research programs.

Strengths and Limitations

As a relatively new area of investigation, multivariate behavior genetic analysis of personality factor structure is in need of further exploration and replication. One limitation of the current study that is important for future research is the number of facets that act as indicators for each trait. By increasing the number of indicators, more complex models can be fit to the data. One such model is a Cholesky model (Loehlin, 1996). This model imposes less structure than the independent pathways model and provides even more strength for a looser interpretation of the coherence of the facets if it fits the data best. We did not report results for this model because it requires at least four indicators in order to be distinguishable from the independent pathways model. A greater number of indicators would also allow one to utilize the approach taken by Jang et al. (2002) to determine the number of coherent genetic and environmental factors that are needed to account for the variance in the facets. They found two factors were needed indicating a lack of genetic unity among the indicators. Again, this method requires several indicators for each trait. Another approach could model genetic and environmental effects as having different structures. Future research that utilizes finer grained analysis of facet level indicators will possess greater flexibility to test these interesting and important possibilities.

A further limitation is the use of a non-standard measure of the FFM. We used a novel scoring algorithm of the CPI which has been shown to produce reliable scores for

the Big Five across several samples (Soto & John, 2009). However, the scale has not been used widely in research due to its recent construction. Additionally, our results do not perfectly match the findings of past research. Johnson and Krueger (2004) utilized a similar methodology in that they specified theoretically meaningful models and classified Neuroticism and Openness differently. While they found that the common pathways model fit Neuroticism and the independent pathways model fit Openness, we found the opposite result. This may have resulted from the different scales utilized in the studies. However, as Costa and McCrae (1992b) point out, "the five-factor model should not be identified with any of its operationalizations; it is an evolving scientific construct, not an instrument" (p. 653), a logic that originally derives from Spearman's (1927) theorem known as "indifference of the indicator." To the extent that the scales used between different studies measure the same latent trait, the differences between them are unimportant. Replication of findings across both different samples and different methods of measurement is much stronger than simple replication across samples (Lykken, 1968).

Another explanation for the differences between this and other studies could be the age of the sample. Johnson and Krueger's (2004) sample was composed primarily of middle aged individuals, and the sample recruited by Jang et al., (2002) was from the general population with a large age range and an average age over 30 years. The current sample was primarily 17 year olds in adolescence. In both of the previous studies, the effect of the shared environment was able to be dropped to model a more parsimonious AE model without significantly worsening fit for all traits, which was not true for the current sample. Many examples and theories exist as to why estimates of quantitative genetic effects would increase with age and shared environmental effects typically decrease. For example, Scarr and McCartney (1983) have hypothesized that as children grow and develop, they are more able and encouraged to create their own environments

rather than live in the same family environment. The data analyzed for this study came from a sample of teenagers who had not yet (or had very recently) moved out of their parents' home whereas in the previous studies, the majority of twins no longer lived together as they were mostly mature adults. The effect that this would have on the distribution of global and specific genetic and environmental effects in addition to the effect on model comparisons is unclear. Future longitudinal work should investigate the multivariate biometric structure of personality at different developmental stages as well as the influence that genes or the environment play in change and stability over time.

Several cautionary notes should also be made of the current findings. This study made use of a fairly large sample of twins. However, multivariate behavioral genetic models must estimate numerous parameters. As such, the finer grained distinctions of parameter estimates should be considered preliminary findings until further replication and confirmation with different and larger samples are obtained. The minor discrepancies between the obtained and expected factor models for the CPI-Big Five also require some caution when drawing conclusions. Adventurousness was the only facet that did not have its primary loading on the intended trait for both twins which may indicate that the sample of relatively high achieving students who took the NMSQT responded differently to items assessing this construct than participants in the original samples (Soto & John, 2009). However, factor and variable congruence coefficients were all very high indicating that the facets utilized in this study are largely the same as found in Soto and John (2009). This demonstrates the usefulness of the new scoring algorithm which opens a large body of previously collected data to be transformed into facets and factors that are relevant in current personality psychology.

In light of the minor discrepancies between the current factor model and that reported by Soto and John (2009), model fit statistics for the biometric model

comparisons may be affected by the lack of coherence in the measurement model, specifically for Adventurousness. In our analyses, the majority of the ACE variation components influencing Adventurousness were found at the specific rather than the common level in both the independent and common pathways models. In the common pathways model that was retained for Openness, the factor loading for Adventurousness was only .23 in comparison to .58 and .66 for Idealism vs. Conformity and Intellectualism, respectively. Keeping in mind that the comparison between the common and independent pathways models only takes into account differences at the common or global levels (all specific effects are modeled exactly the same in both model types), it makes sense that Openness would be less likely to display significant misfit when fitting the constrained model. The global effects are primarily dependent on the relationship between two facets with Adventurousness not being strongly related at the global level. Thus, constraining the three facets to have a common structure produces little model misfit because the majority of the variation for Adventurousness is found at the specific level for both the common and the independent pathways models. As such, the results regarding Openness in this study may be sample specific and should be interpreted with caution.

The example of Openness proposes a further limitation and strength of this study. As mentioned, the model comparisons performed in the study only took into account differences in model misfit of global effects. Models were specified in this particular way to mimic the hypothesized hierarchical structure of personality traits. However, judging that Openness is best represented by the common pathways model when one of the facets was largely independent of the other two facets presents a large problem that calls for careful thinking and interpretation rather than strict reliance on statistical tests. In this instance, the preliminary descriptive goal of the present study to explore the general

structure of the genetic and environmental effects with a relatively unconstrained model becomes a crucial aid to interpretation. Further, the analysis of where misfit occurs for each of the Big Five provides further information about the ability to structure specific effects as common or independent. The central finding of the current study is that there are significant facet-specific, genetic effects that are overshadowed when aggregated to the factor level.

Finally, it is of note that all models fit for the current project have good RMSEA values, indicating that both the common and independent pathways models were adequate approximations of the observed data. While the χ^2 difference tests that we used to compare models are sensitive to absolute fit, RMSEA is an approximate fit index. The results of the model comparisons can be interpreted as differences in exact fit to the data even though both models fit reasonably well. However, low RMSEA values for the common pathways model may not have persisted had more indicators for each factor been available. As the number of indicators increases, the differences between the common and independent pathways models also increase in terms of degrees of freedom and the amount of constraint placed on the facet structure. Additionally, one might argue that our χ^2 difference tests were somewhat overpowered, in light of the good overall fit of both model types. While it is accurate that the χ^2 difference test is sensitive to sample size, in the current study the test determined that the more parsimonious common pathways model did not fit significantly worse than the independent pathways model for two traits. Had the sample size been so large as to overpower the χ^2 difference test, one would expect to reject the common pathways model in all instances. In light of the fact that the model fit statistics were relatively good for both models, it would be more appropriate to view the difference test as shedding light on whether or not common genetic and environmental effects can be constrained to be proportional across facets

instead of proving one model "right" or "wrong" since both are largely "right," in that they approximate the observed patterns of relations in the raw data adequately. The results of the current study would imply that the common genetic and environmental effects of facets for Extraversion and Openness act in a more uniform manner than the common effects of Agreeableness, Conscientiousness, and Neuroticism. Of primary importance, all facets were found to display specific variation unique of the other facets within a trait, with some displaying substantial genetic variation not shared with the higher order trait. Overall, these results indicate that a complete understanding of human personality will require research spanning multiple levels of both description and explanation.

Appendix

The independent pathways model specifies that each facet has separate genetic, shared environmental, and nonshared environmental loadings on common (domain-general) genetic, shared environmental, and nonshared environmental factors. In comparison, the common pathways model specifies that each facet has a single phenotypic loading on a common phenotypic factor, which in turn has loadings on a single set of trait-level genetic, shared environmental, and nonshared environmental factors. The following parameter constraints can be applied to the common factor loadings from the independent pathways model to function as a common pathways model:

$$\lambda \begin{bmatrix} \boldsymbol{A_f} \\ \boldsymbol{C_f} \\ \boldsymbol{E_f} \end{bmatrix} = \lambda_f \times \lambda \begin{bmatrix} \boldsymbol{A_t} \\ \boldsymbol{C_t} \\ \boldsymbol{E_t} \end{bmatrix},$$

where λ represents common factor loadings, the subscript f indicates that a parameter is allowed to vary across facets, and the subscript t indicates that a parameter is

constant across all facets belonging to the same trait. Note that facet-specific genetic and environmental loadings are not represented above; $\lambda[A_f]$, $\lambda[C_f]$, and $\lambda[E_f]$ only represent loadings on the common genetic and environmental variance components.

To illustrate how these constraints function, we can apply the above constraints to the facets of

Agreeableness yielding:

$$\lambda \begin{bmatrix} \boldsymbol{A_{trust}} \\ \boldsymbol{C_{trust}} \\ \boldsymbol{E_{trust}} \end{bmatrix} = 1 \times \lambda \begin{bmatrix} \boldsymbol{A_{Agreeablness}} \\ \boldsymbol{C_{Agreeableness}} \\ \boldsymbol{E_{Agreeableness}} \end{bmatrix},$$

$$\lambda \begin{bmatrix} \boldsymbol{A_{compassion}} \\ \boldsymbol{C_{compassion}} \\ \boldsymbol{E_{compassion}} \end{bmatrix} = \lambda_{\text{compassion}} \times \lambda \begin{bmatrix} \boldsymbol{A_{Agreeablness}} \\ \boldsymbol{C_{Agreeableness}} \\ \boldsymbol{E_{Agreeableness}} \end{bmatrix}, \text{ and}$$

$$\lambda \begin{bmatrix} \boldsymbol{A_{humility}} \\ \boldsymbol{C_{humility}} \\ \boldsymbol{E_{humility}} \end{bmatrix} = \lambda_{\text{humility}} \times \lambda \begin{bmatrix} \boldsymbol{A_{Agreeablness}} \\ \boldsymbol{C_{Agreeableness}} \\ \boldsymbol{E_{Agreeableness}} \end{bmatrix}.$$

Note that the first phenotypic loading (what would ordinarily be λ_{trust}) is fixed to 1 to set the metric of the latent phenotype. It can be seen that without the constraints on the right hand side of the equations, there are 9 free parameters that would be individually estimated, whereas with the constraints, 5 free parameters are estimated (three trait-level biometric components and k-1 phenotypic loadings, where k is equal to the number of facets).

Chapter 2: Genetic and Environmental Continuity in Personality Development – A Meta-Analysis

A defining feature of psychological traits is that they are relatively stable across time (Allport, 1937). Differential stability -- the stability of an individual's rank order over time -- is typically assessed with a test-retest correlation and is central to understanding the function of traits over time. A great deal of research on the differential stability of personality traits supports their classification among the most stable psychological traits (Conley, 1984). Nevertheless, personality is not uniformly stable over development. For instance, an influential meta-analysis of test-retest data from 152 longitudinal 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 (2000, p. 4) 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."

The current project meta-analyzes longitudinal, genetically informative data to address this issue. We ask three core questions: (1) What are the relative contributions of genes and the environment to personality across the lifespan? (2) How stable are the genetic and environmental influences on personality across the lifespan? (3) To what extent do genetic and environmental factors contribute to the increasing stability of personality across the lifespan? We begin by reviewing influential theoretical models of personality development and discuss how behavioral genetic methodology can be used to 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 psycho-sexual development (Freud, 1959/1908)

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 (McAdams & Olsen, 2010). Erikson (1963/1950) saw personality as developing in relation to success or failures with certain social challenges. Similarly, attachment theories have posited that early caregiving experiences have a lasting influence on a multitude of aspects of an individual's life (Bowlby, 1964).

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 neo-socioanalytic (Roberts & Woods, 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 lifespan 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, 2005, 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 "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. Empirical evidence for this perspective comes from a number of sources beyond that claimed by McCrae et al. (2000) including molecular genetic findings supportive of genetic influences on personality (de Moor et al., 2010; Terracciano et al., 2010; Webb et al. 2012; Vinkhuyzen et al., 2012), parallel cross-cultural age trends in average trait levels (McCrae et al., 1999), 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, DeRubeis, Hollon, Amsterdam, Shelton, & Schalet, 2009), to name only a few examples.

Based on early work examining the stability of personality across the lifespan, 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 lifespan 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 this 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, an outcome 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 re-order 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 re-ordered in terms of their height. McCrae and Costa (2006) argue that personality traits can be understood in a similar way. Ultimately, Costa and McCrae (2008) claim "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, it

would predict that environmental influences are largely unstable and unimportant for increasing phenotypic stability.[1]

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 2000's, Roberts and colleagues (Caspi & Roberts, 2001; Caspi, Roberts, & Shiner, 2005; Roberts & Wood, 2006; Roberts, Wood, & Caspi, 2008) have laid out a growing number of postulated principles and mechanisms for social maturation (there are currently 26[2]) 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 emphases the causal nature of socialization.

[1] In the most recent, complete exposition of the FFT, Costa and McCrae (2008, p. 163) include 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 Costa and McCrae (2008, p. 168) argue that this effect is "outside the confines of personality proper." Therefore, a charitable interpretation of the FFT would be that some environmental effects may exist, but that major lifespan trends should still be predominantly driven by genetic factors.

[2] This number is based on a count taken from Roberts et al. (2008).

Most relevant to the current discussion are Roberts and colleagues' "cumulative continuity principle" and the "plasticity principle." These principles hold that personality displays increasing stability throughout the lifespan (past age-30) and that personality is always amenable to change. According to this perspective, personality stability continues to increase throughout the lifespan 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 et al., 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, Lüdtke, & Trautwein, 2012), negative life events (Kandler, Bleidorn, Riemann,

Angleitner, & Spinath, 2012), macro-environmental differences (Bleidorn, Kilmstra, Denissen, Rentfrow, Potter, & Gosling, 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) has 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 (2009) pointed to consistent, bottom-up effects of the environment onto 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" (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 the continued approval of students for this type of behavior. The student approval generates satisfaction in a social role which 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 lifespan, 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 lifespan. Increases in phenotypic stability would largely be explained by increasingly stable contributions from the environment with genetic contributions remaining largely constant across the lifespan.

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 genetic influences are activated or inactivated in response to new environmental situations, such as entry into school (Johnson, 2007; Krueger, Johnson, & Kling, 2006). Age-related activation of genetic effects can also result from biological changes associated with development, such as puberty. Moreover, dynamic transactional processes of *gene-environment correlation* occur when individuals select and evoke different environments on the basis of their genetically influenced preferences, motivations, and traits, which in

turn affect psychological development (Plomin, DeFries, & Loehlin, 1977; Scarr & McCartney, 1983).

Empirical evidence for gene-environment interaction and gene environment correlation comes from both molecular genetic and quantitative genetic studies. For instance, with respect to gene-environment interaction, researchers have reported differential associations between parenting received and personality for individuals with different variants of 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). Research using twin and family methods have found that the magnitudes of genetic influences on personality differ across environmental contexts (Ganiban, Ulbricht, Saudino, Reiss, & Neiderhiser, 2011; Krueger, South, Johnson, & Iacono, 2008). 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 in that siblings that are more genetically similar are more likely to experience more similar environments (Avinun & Knafo, 2013; Kendler & Baker, 2007; Riemann, Kandler, & Bleidorn, 2012; Sturaro, Denissen, van Aken, & Asendorpf, 2008; Saudino & Plomin, 1997). As Plomin, Reiss, Hetherington, and Howe (1994, p. 32) have 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 towards engaging with different sorts of environments. For example, McAdams, Gregory, and Eley (2013) found that the genes that influence exposure to negative life events and parenting could be accounted for by the genes that influence an individual's personality characteristics.

Both gene-environment correlation and gene-environment interaction have implications for personality development and provide insights into the types of lifespan trends to expect for behavioral genetic outcomes. Early in the lifespan, children may differentially respond to the common environment provided by their caregivers on the basis of genotype. This type of gene-by-environment interaction results in environmental effects becoming coupled with genetic variation. Genetically identical individuals would respond to the environment similarly and become *more* similar than individuals that share fewer genes. If the same sorts of environments either recur over development or if early environments have a lasting impact, then genetic stability will be high. 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-by-environment interactions for personality development are fleeting and give way to less environmentally dependent gene expression (e.g., Conley, 1984; Wolf & Weissing, 2012). If this is the case, then it would be expected that heritability would decline with age, and that genetic stability would be low in early life and high later in the development, once the phenotype has stabilized. In adolescence and adulthood, when individuals leave their home environments, it is more likely that the unique environment that an individual is located in will have an impact on personality rather than the common family environment

(Bouchard & Loehlin, 2001). Gene-by-environment interaction in this context has the effect of making genetically identical individuals *less* similar, and therefore genetic effects become tied to environmental variation. In this case, environmentality would be expected to rise as individuals mature and leave the shared home environment. Again, such an interaction may be recurring, lasting, or fleeting. If it is recurring or lasting, this would result in stable environmental influences, and fleeting interactions would cause instability of environmental influences.

Active and evocative gene-environment correlation can result in true environmental effects being experienced nonrandomly, such that they are tied to genotypic differences. 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, all else being equal, strong patterns of gene-environment correlation would be expected to manifest in increasing heritability and 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, whereas 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. The stability of nonshared environmental effects on personality, then, might be expected to be relatively lower than that of genetic effects on personality.

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, behavior genetic research also provides insight into the stability of genetic and environmental influences, and the relative contribution of genetic and environmental mechanisms to trait stability. In this section we provide an overview of both univariate (in which a single construct is measured at a single time point) and longitudinal behavioral genetic models, and formally define what we mean by stability of genetic effects, stability of environmental effects, and genetic and environmental contributions to stability.

Classical behavior 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 do dizygotic twins. Variance in the trait is formally decomposed into an additive genetic component (A) 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). A third component, termed the shared environment (i.e., experiences that causes siblings raised within the same home to be more similar to each other, typically abbreviated C) can be derived, but 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 is presented in Figure 3. 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 example, for monozygotic twins (who have 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, and 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 they are always uncorrelated no matter what sibling type is represented. When 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.

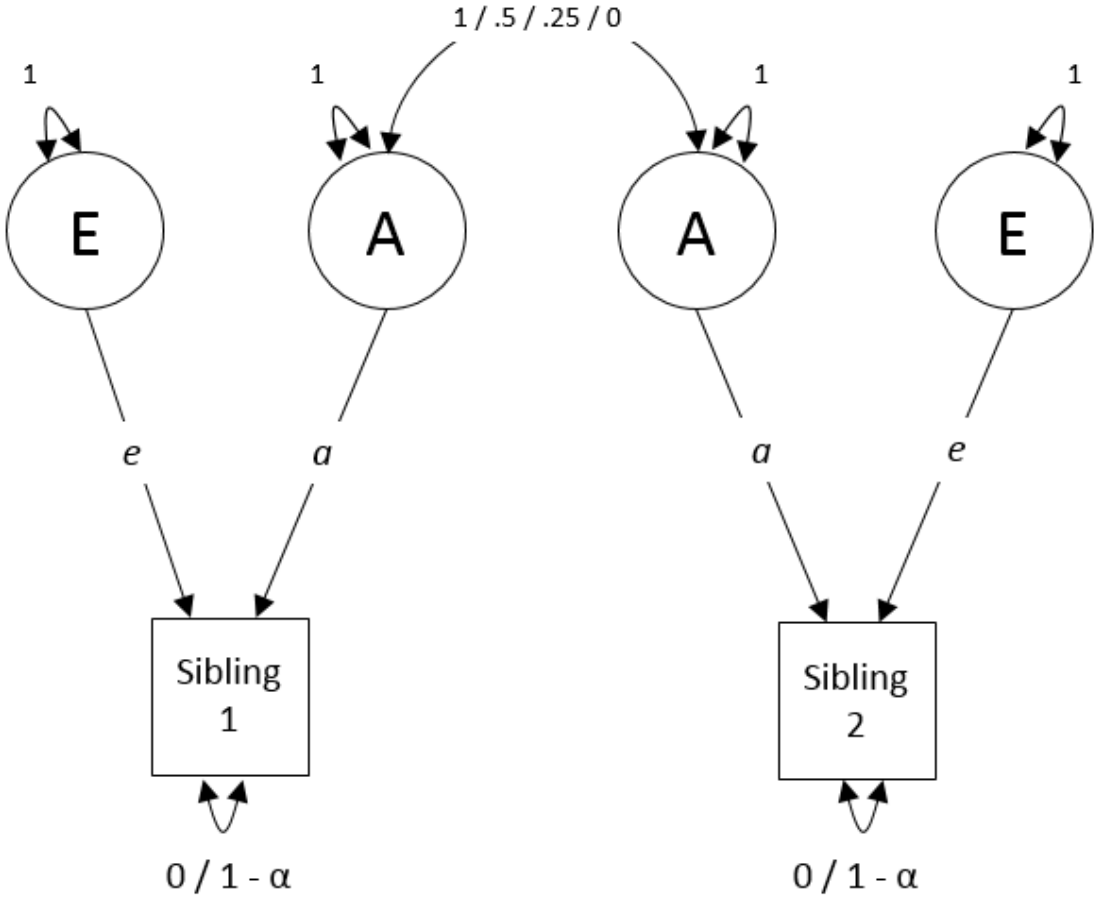

Figure 3. Univariate behavior genetic model that decomposes variance in a trait (indicated as a square) into that which is due to latent genetic (A) and environmental (E) components (indicated as circles). The correlation between genetic factors is specified for each group depending on the known genetic 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 – 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.

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 correlated factors model presented in Figure 4 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 it is the same or different genetic and environmental effects 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 would be indicative of shared genetic influence across the 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 sib. 1 is only correlated with E at time 2 for sib. 1). Genetic effects, however, are correlated across siblings and time points to the extent that the siblings share genetic material, and this 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 × rA), but for dizygotic twins, this pathway would be constrained to be equal to half of the genetic correlation (rA* = 0.5 × rA). Because the same individual retains the same genotype across time points, the within-sibling cross-time pathway is equal to rA.

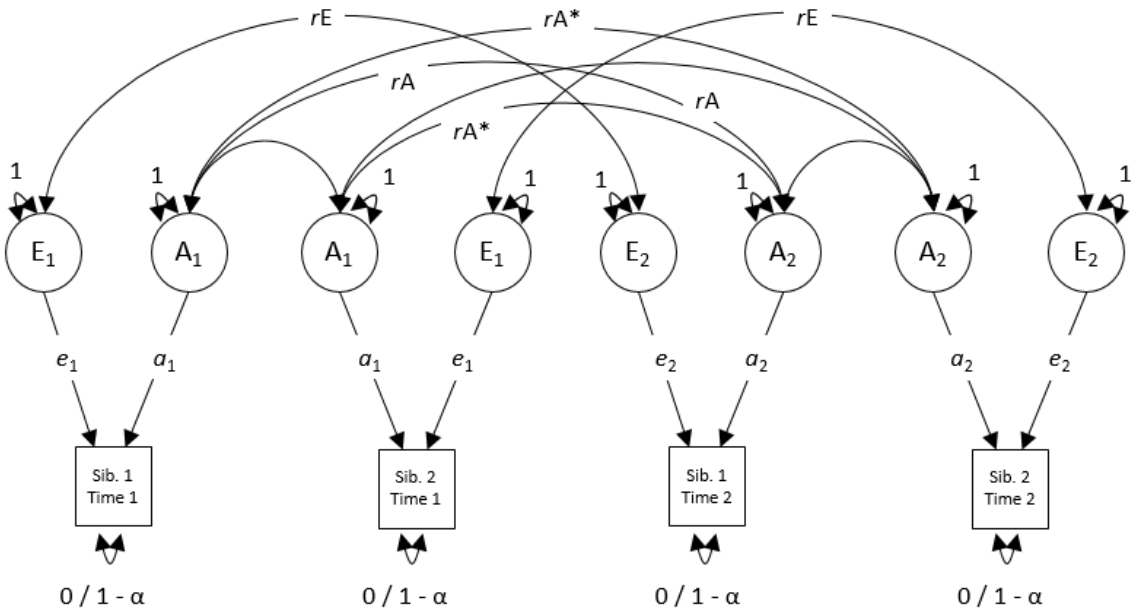

Figure 4. Longitudinal correlated factors model that decomposes variance in repeated assessments of personality into that due to time-specific genetic (A) and environmental (E) components, as well as the temporal stability of the genetic (*r*A) and environmental (*r*E) components. Interpretation of the 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 *r*A. Cross-sibling cross-time environmental influences are constrained to zero by definition.

Importantly, genetic stability (*r*A) is not directly tied to the magnitude of genetic influence on the trait (*a*). An outcome that is highly heritable may be influenced by the same genes at two time points, and similarly, an outcome that is largely influenced by the

environment may still be influenced by a small, but stable, set of genes. Likewise, an outcome may reflect minimal environmental effects, but if these influences are stable across time, then environmental stability will be high. Understanding the genetic contribution to observed stability requires information concerning stability and heritability. The contribution of genetic and environmental effects to stability is calculated by multiplying the time-specific variance pathways by the associated latent variable correlation. For example, $a_1 \times r\text{A} \times a_2$ yields the genetic contribution to stability. Summing the contribution from A and E recaptures the standard test-retest correlation (e.g., phenotypic stability = $a_1 \times r\text{A} \times a_2 + e_1 \times r\text{E} \times e_2$). Thus, if an outcome is highly heritable and these genetic influences are largely stable across time, then genes will contribute substantially to stability. However, genetic influences may change over time even if heritability is high at both time points. In such a case genes would not contribute much to stability. Similarly if genes are highly stable but heritability is low at both time points, genes would not contribute much to stability. If an environmental influence is particularly lasting, such as mature romantic or occupational roles, then environmental stability will be high. However, if environmental influences primarily affect temporary states, then environmentality may be high, but environmental stability would be low. In this situation, the environment would not account for differential stability.

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 genes to phenotypic stability would be .40 (i.e., $\sqrt{.5} \times .8 \times \sqrt{.5}$), and the bivariate heritability would be .67 (i.e., $[\sqrt{.5} \times .8 \times \sqrt{.5}]/.6$). Thus, 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* 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) has recently 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. This attenuates the true amount of stable variation in personality. Ferguson's (2010) 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 age 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.

Behavior 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 zero 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 3 and 4 explicitly depict this by showing that the phenotype's residual variance is zero in the standard model (this is often left out of depictions of behavior genetic models) and as 1-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 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 *r*E to indicate the amount of true environmental stability removed of idiosyncratic measurement error. However, the contribution of the environment to phenotypic stability (i.e., $e_1 \times r\text{E} \times e_2$) would remain unchanged, and therefore, the estimated phenotypic stability would also remain unchanged. Because increases in *r*E are counterweighted by decreases in e_1 and e_2, correcting for measurement error does not affect aggregates of these parameters. With this 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 disattenuation based on Cronbach's α can produce illogical estimates of "true" correlations because α overestimates measurement error under several circumstances. In the current context, this can have the effect of underestimating environmentality and overestimating environmental stability. We will 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. This is because it is nearly a

truism that traits are less stable over longer periods of time. Fraley and Roberts (2005) specifically emphasized this point by re-analyzing 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, will control for the linear effect of time lag in all analyses. However, the emphasis of the current meta-analysis is on lifespan age-based trends. In order to confine the scope of this article, we will 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 will examine whether different traits reflect different developmental patterns in terms of the stability of genetic and environmental effects.

Self- vs. 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 towards 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 implies that self-reports of personality may be more likely to show evidence of increasing stability with age. In spite of this, previous work has found little evidence that self- and peer-reports differ for behavioral genetic analyses (Riemann, Angleitner, & Strelau, 1997).

Broad vs. 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 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 smoothed population-level age trend in differential stability in spite of discontinuities in individual-level developmental trajectories. For the current meta-analysis we will 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 meaningfully important in the literature, and a continuous parametric approach which represents development more progressively. Importantly, 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 will model turning points for each of the developmental trends at the ages of 3 years (temperament to traits)[3], 15 years (childhood through 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 or action across situations or time (McCrae et al., 2000). Empirically, temperament has been consistently linked with trait-like measures (Caspi, Harrington, Milne, Amell, Theodore, & Moffitt, 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 process (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 likely view age 30 as an important turning point and the others to be relatively inconsequential. Conversely, the neo-analytic and sociogenomic models of personality would be more

[3] In the studies meta-analyzed for this article, age 3 years was the last age with which temperament was specifically measured in young children.

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, behavior genetic studies of personality development. This is the first comprehensive empirical meta-analysis of this literature. The only previous synthesis of the literature is by Kandler (2012). In a short review of the longitudinal, behavior 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, 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: (1) the levels of heritability and environmentality of traits at one point in time; (2) the test-retest stability of phenotypic traits and of genetic and environmental effects; and (3) the contribution of genetic and environmental effects to test-retest stability. Based on our review above, the FFT would predict that genetic effects are large (i.e., high heritability), stable (i.e., high *r*A), 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 *r*E), 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., non-trivial estimates of heritability, *r*A, and genetic contribution), but increases in phenotypic stability would be primarily mediated by increasingly stable environmental factors (i.e., increasing environmentality and *r*E with age).

METHOD

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 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 behavior 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 trends 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, Donnellan, Blonigen, Krueger, McGue, Iacono, & Burt, 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 9 presents the citation, dataset, 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.[4]

[4] 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 behavior genetic models such as the cholesky decomposition, the correlated factors model and the simplex model, as well as variations of these.

Table 9. Description of included studies

Citation	Dataset	Variables Examined	Number of Waves	Age Range (years)	Longitudinal Sample Size (pairs)
Blonigen et al. (2008)[a] *	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 et al. (2006)[a] *	Minnesota Twin Family Study	Fearless Dominance; Impulsive Antisociality	2	17.00-24.00	360 MZT; 188 DZT
Bratko & Butkovic (2007) *	Zagreb Community Sample	Extraversion; Neuroticism; Psychoticism; Lie Scale	2	17.00-21.00	75 MZT; 85 DZT
De Fruyt et al. (2006) **	Small-scale twin family study	Emotional Instability; Extraversion; Imagination; Benevolence; Conscientiousness	2	8.65-11.65	79 MZT; 124 DZT
Ericson et al. (2011)[b] *	Southern California Twin Project	Cognitive-Perceptual; Disorganization; Interpersonal-Affective	2	11.89-14.69	205 MZT; 293 DZT
Forsman et al. (2008) *	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) **	Wisconsin Community Sample	Anger	2	1.00-3.00	130 MZT; 237 DZT
Ganiban et al. (2008) **	Nonshared Environment and Adolescent Development	Activity; Emotionality; Shyness; Sociability	2	12.80-15.45	63 MZT; 228 DZT; 60 HST; 44 URT
Gillespie et al. (2004) *	Brisbane Community Sample	Extraversion; Neuroticism; Psychoticism	3	12.00-16.00	216 MZT; 192 DZT
Hopwood et al. (2011)[a] *	Minnesota Twin Family Study	Constraint; Negative Emotionality; Positive Emotionality	3	17.00-29.00	349 MZT; 183 DZT

Table 9, cont.

Johnson et al. (2005) *	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) ***	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)	156 MZT; 95 DZT (young cohort) 140 MZT; 88 DZT (middle cohort)
Kupper et al. (2011) *	Netherlands Twin Registry	Negative Affectivity; Social Inhibition	3	17.20-29.50	650 MZT; 650 DZT
McGue et al. (1993) *	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
Niv et al. (2011)[b] *	Southern California Twin Project	Inattention-Impulsivity; Motor Impulsivity; Non-planning Impulsivity	2	11.89-14.69	205 MZT; 293 DZT
Pedersen & Reynolds (1998) *	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 et al. (2006) *	OCTO-Twin Study	Extraversion; Neuroticism	3	82.30-86.30	149 MZT; 202 DZT
Rietveld et al. (2004) **	Netherlands Twin Registry	Attention Problems	4	3.00-12.00	1891 MZT; 3310 DZT

Table 9, cont.

Saudino (2012) **	Boston University Twin Project	Activity Level	2	2.00-3.00	141 MZT; 163 DZT
Saudino & Cherny (2001)† **	MacArthur Longitudinal Twin Study	Activity; Affect-Extraversion; Behavioral Inhibition; Shyness; Task Orientation	4	1.17-3 .00	118 MZT; 106 DZT
Spengler et al. (2012) *	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) *	Keio Twin Project	Behavioral Inhibition Scale; Behavioral Activation Scale	2	23.27-25.52	85 MZT; 32 DZT
Viken et al. (1994) *	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; 1572 DZT 580 MZT; 1180 DZT 405 MZT; 798 DZT 288 MZT; 618 DZT 216 MZT; 498 DZT 171 MZT; 339 DZT
Zavos et al. (2012) *	The G1219	Anxiety Sensitivity	3	15.00-17.00	345 MZT; 895 DZT

Note. Citations marked with the same superscript letter use the same data. MZT refers to monozygotic twins reared together; DZT refers to dizygotic twins reared together; MZA refers to monozygotic twins reared apart; DZA refers to dizygotic twins reared apart; HST refers to half-siblings reared together; URT refers to unrelated siblings reared together.
† Data on Affect-Extraversion was not available for wave 4, and thus only data on ages 1.17 to 2 years were analyzed.
* indicates self-report format
** indicates informant-report format
*** indicates both formats

We used M*plus* statistical software (Muthén & Muthén, 1998-2010) to fit a longitudinal correlated factors model (Figure 4, 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 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 are 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 4. 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), and therefore the reported results may be considered lower-bound estimates of precision.

Having completed this process, we compiled each parameter depicted in Figure 4 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 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 = .20). As these estimates tended to carry comparatively small weight in the analyses, we chose to leave them in the dataset as estimated. The patterns of results are unchanged if these estimates are removed or censored to the closest reasonable estimate.

We formed two datasets: one arranged in a cross-sequential manner with heritability and environmentality at each time point, and another with the phenotypic stability, genetic and environmental correlations and contributions to phenotypic stability associated with each *pair* of time points. This resulted in 330 × 3 (heritability, environmentality, and corrected environmentality) effect sizes for the cross-sequential dataset and 251 × 6 (phenotypic stability, genetic stability, and corrected environmental stability, and the contributions of genes and the environment to phenotypic stability) effect sizes in the longitudinal dataset. In the cross-sequential dataset, each study contributed an average of 28.66 sets of effect sizes (SD = 17.74, range = 2-60). In the longitudinal dataset, 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 dataset and age at the initial time point and the time interval between measurements in the second dataset. Additional variables included in the dataset are described below.

Study/Variable Characteristics

Error Correction

For most articles (k = 16), we were able to obtain calculated reliability estimates (Cronbach's α) within the sample at each time point and for each measure which is 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). While 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) reports 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 (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 (Tellegen & Waller, 2008), the Neuroticism-Extraversion-Openness Five-Factor Inventory (Costa & McCrae, 1989), the Neuroticism-Extraversion-Openness Personality Inventory – Revised (Costa & McCrae, 1992), the Behavioral Inhibition Scale/Behavioral Activation Scale (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), 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). 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 non-specified traits (coded as 0). Extraversion was the most studied trait (n = 98 across all time points)[5], followed by neuroticism (n = 94), conscientiousness (n = 67), agreeableness (n = 43), and openness to experience (n = 30).

Self- vs. 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 (n = 253), with a sizeable minority using other-report (n = 77).

Broad vs. 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 Multidimensional Personality Questionnaire (e.g., constraint, negative emotionality, positive emotionality; 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] All n's reported for the moderators refer to specific time points rather than pairs of measurements.

.5 and broad measures as .5. Slightly more measures were classified as narrow (n = 170) than broad (n = 160).

Analytic Approach

Our primary goal was to test a series of alternative models that specified different lifespan trends for the genetic and environmental influences on personality development. To accomplish this, 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 continuous trend or display important deviations from the continuous trend.[6]

[6] 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

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 M*plus* 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: (1) age-trends in the level of heritability, environmentality, and corrected environmentality across the lifespan; (2) age-trends in stability at the phenotypic, genetic, environmental, and corrected environmental level across the lifespan; and (3) age-trends in the contribution of genetic and environmental effects to phenotypic stability across the lifespan. For this 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 is 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

---

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.

interpretation and graphical presentation of the age-related effects assumes 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

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

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), Akaike's Information Criterion (AIC), and Bayesian Information Criterion (AIC). 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 this represents a more parsimonious account of the data. 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.

A final goal was to determine if any of these trends 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- vs. other-report, and

broad vs. 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 10 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 sizeable standard deviations. The studies covered nearly the entire range of the lifespan 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 composition of the samples was not well-reported in the articles. Based on 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 10 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 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.

Table 10. Characteristics of the Studies Meta-Analyzed

Variable	Mean	SD	Range
Age at baseline (years)	23.02	19.99	1.00-84.30
Interval (years)	5.43	2.91	.33-13.80
Age at follow-up (years)	28.45	20.53	1.67-86.30
Percent Female	54.12	9.46	.00-100
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.

Age-Trends in the level of Heritability and Environmentality

Table 11 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 12 presents the results of the fixed-effects, exponential models for the outcomes, and Figure 5 presents the raw data and best fitting models graphically. In Figure 5, 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.

Table 11. Age trends for heritability, environmentality, and corrected environmentality

Regression Model	b_0	$b_{0\text{-}3}$	$b_{3\text{-}15}$	$b_{15\text{-}30}$	$b_{30\text{-}60}$	$b_{60\text{-}90}$	τ	LL	AIC	BIC
Heritability										
Mean Effect Size	.480 (.020) ***	-	-	-	-	-	.092 (.016) ***	-390.508	785.016	792.614
Full Model	.547 (.129) ***	.041 (.047)	-.015 (.005) **	-.004 (.003)	.000 (.001)	-.002 (.001) *	.061 (.012) ***	-372.163	758.326	784.920
Preferred Model (AIC)	.541 (.131) ***	.044 (.049)	-.017 (.005) ***	-.001 (.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
Environmentality										
Mean Effect Size	.515 (.020) ***	-	-	-	-	-	.091 (.016) ***	-395.508	795.016	802.614
Full Model	.453 (.128) ***	-.041 (.047)	.014 (.006) **	.005 (.003)	.000 (.001)	.002 (.001) *	.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

Table 11, cont.

Corrected Environmentality									
Mean Effect Size	.311 (.023) ***	-	-	-	-	-	.083 (.017) ***	-388.250	780.499	788.085
Full Model	.307 (.194)	-.042 (.077)	.009 (.006)	.006 (.004)	.000 (.001)	-.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. b_0 represents the intercept; $b_{0\text{-}3}$ represents the coefficient for the first linear segment, $b_{3\text{-}15}$ represents the following segment and so on until $b_{60\text{-}90}$. For simplified models, arrows are used to indicate the parameter estimate that applies during the given developmental period. τ represents the random-effect standard deviation in the between study effect size.

* indicates $p < .05$; ** indicates $p < .01$; *** indicates $p < .001$.

Table 12. Results of continuous exponential models for each developmental outcome

Outcome	b_0	b_1	b_2	$b_{\text{time lag}}$	LL	AIC	BIC
Heritability	.336	-.405	-.037				
	(.074)	(.068)	(.024)	-	-418.390	2916.778	2928.175
	***	***					
Environmentality	.665	.406	-.037				
	(.074)	(.068)	(.023)	-	-421.816	2925.297	2936.694
	***	***					
Corrected	.451	.388	-.033				
Environmentality	(.074)	(.089)	(.020)	-	-1493.071	2997.071	3016.067
	***	***					
Phenotypic	.706	.349	-.055	-.006			
Stability	(.048)	(.075)	(.022)	(.006)	-414.479	4944.914	4959.016
	***	***	*				
Genetic Stability	.997	.506	-.070	-.011			
	(.016)	(.040)	(.015)	(.004)	-290.729	4405.366	4419.468
	***	***	***	**			
Environmental	.539	.403	-.080	-.025			
Stability	(.779)	(.594)	(.808)	(.012)	-393.141	4421.708	4435.810
				*			
Corrected	1.014	.767	-.029	-.017			
Environmental	(.168)	(.132)	(.021)	(.010)	-357.949	3845.433	3859.519
Stability	***	***					
Genetic	.382	.356	-.860	-.003			
Contribution to	(.029)	(.140)	(.210)	(.007)	-290.733	4321.622	4335.707
Stability	***	*	***				
Environmental	.330	.306	-.045	-.007			
Contribution to	(.044)	(.039)	(.018)	(.003)	-281.284	4798.369	4812.455
Stability	***	***	*	*			

Note. * indicates $p < .05$; ** indicates $p < .01$; *** indicates $p < .001$.

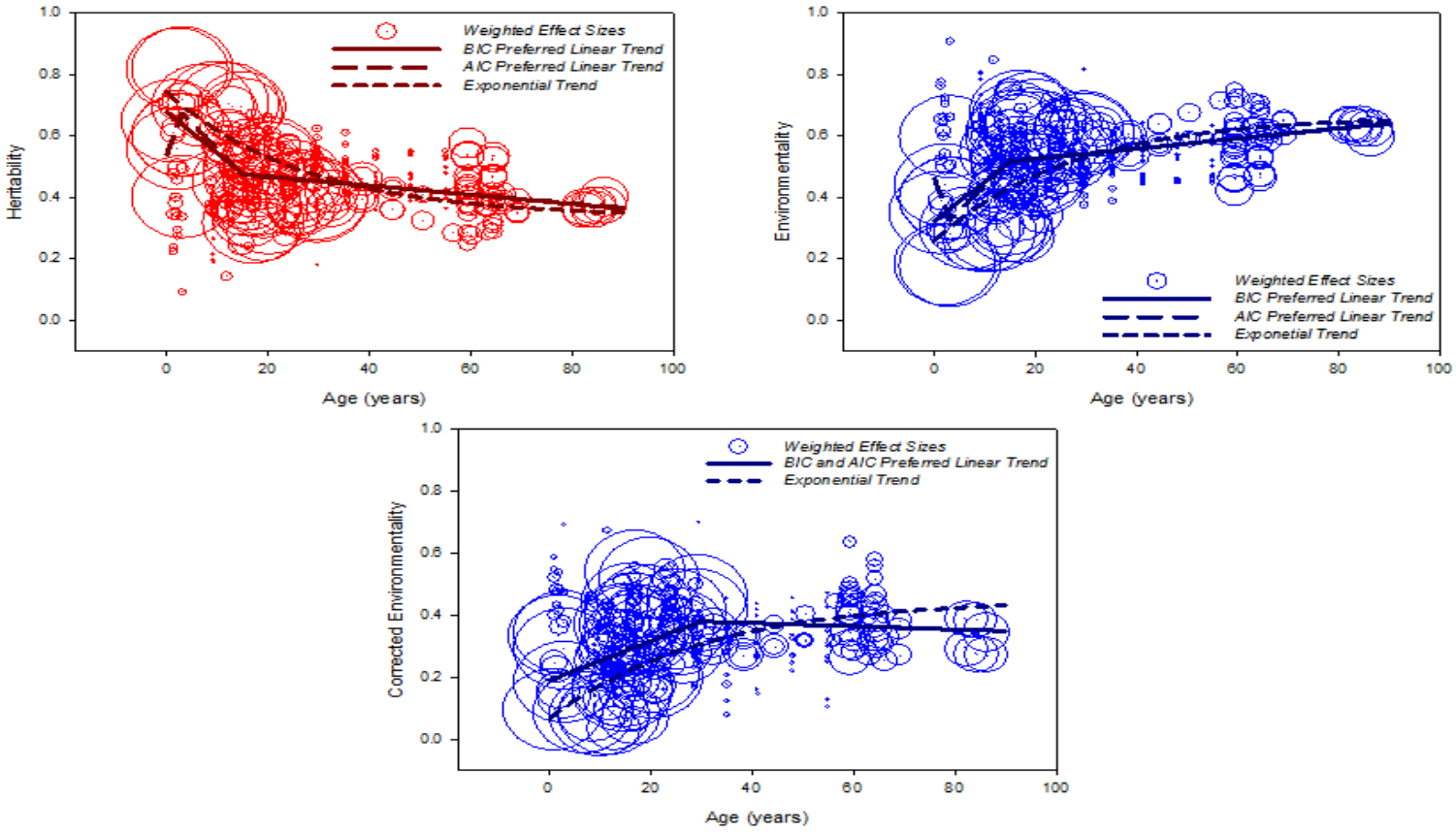

Figure 5. Age-trends in heritability, environmentality, and measurement error corrected environmentality. Circles surrounding data points are scaled by the weighting variable (described in Analytic Approach section) such that larger circles carried more weight in the analysis.

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 lifespan. The full connected-linear model for heritability indicates a substantial influence of genes 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 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 $b_{0\text{-}3}$ column pointing towards the parameter estimate in the $b_{3\text{-}15}$ column indicating that this slope applies in the age range of 0 to 3. Similarly, arrows pointing towards the earlier time point are presented for columns $b_{30\text{-}60}$ and $b_{60\text{-}90}$ indicating that the $b_{15\text{-}30}$ slope applies during this area of the lifespan. Because parameter estimates are adjacent in the $b_{3\text{-}15}$ and $b_{15\text{-}30}$ 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 will 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 ($b_{0\text{-}15}$ = -.013, $p < .01$) in magnitude while environmental effects increase ($b_{0\text{-}15}$ = .013, $p < .05$) across early childhood. Following age 15, relatively small, but significant, changes occur in genetic ($b_{15\text{-}90}$ = -.001, $p < .05$) and environmental ($b_{15\text{-}90}$ = .002, $p < .05$) effects. This 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 ($b_{0\text{-}30}$ = .006, $p < .01$) from accounting for roughly 20% of the variance in early childhood and remains stable after age 30 ($b_{30\text{-}90}$ = -.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 5. 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 lifespan. However, the different models are highly similar, overlap for much of the lifespan, and each visually reflect 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, 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 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. This 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 the Phenotypic, Genetic, and Environmental Stability

Table 13 presents the connected-linear model results for genetic and environmental stability as predicted by age at baseline and time interval between measurements. Table 12 presents the results of the continuous exponential model, and Figure 6 presents the preferred trend lines and meta-analytic data graphically.

Table 13. Age trends for phenotypic, genetic, environmental, and corrected environmental stability

Regression Model	b_0	$b_{0\text{-}3}$	$b_{3\text{-}15}$	$b_{15\text{-}30}$	$b_{30\text{-}60}$	$b_{60\text{-}90}$	$b_{\text{time lag}}$	τ	LL	AIC	BIC
Phenotypic Stability											
Mean Effect Size	.554							.100			
	(.036)	-	-	-	-	-	-	(.016)	-403.240	810.479	817.530
	***							***			
Full Model	.093	.126	.001	.016	.000	-.003	-.006	.065			
	(.123)	(.051)	(.004)	(.003)	(.002)	(.002)	(.004)	(.018)	-334.179	684.358	712.561
		*		***				***			
Preferred Model (AIC	.092	.126	.000	.017	-.001		-.006	.065			
and BIC)	(.128)	(.051)	(.004)	(.002)	(.001)	←	(.004)	(.018)	-334.513	683.025	707.703
		*		***				***			
Genetic Stability											
Mean Effect Size	.844							.119			
	(.038)	-	-	-	-	-	-	(.021)	-340.007	684.014	691.065
	***							***			
Full Model	.518	.030	.014	.017	-.002	-.001	-.013	.036			
	(.123)	(.043)	(.005)	(.003)	(.001)	(.001)	(.006)	(.010)	-287.580	591.159	619.363
	***		**	***	*		*	***			
Preferred Model (AIC	.553			.016	-.001		-.013	.034			
and BIC)	(.024)	→	→	(.001)	(.001)	←	(.006)	(.009)	-287.779	585.558	603.185
	***			***			*	***			
Environmental Stability											
Mean Effect Size	.397							.136			
	(.041)	-	-	-	-	-	-	(.023)	-364.406	732.813	739.864
	***							***			
Full Model	-.275	.189	.003	.010	.004	-.007	-.016	.086			
	(.146)	(.060)	(.006)	(.004)	(.002)	(.002)	(.007)	(.020)	-315.639	647.279	675.482
		**		**	**	***	*	***			
Preferred Model (AIC	-.253										
and BIC)	(.143)	.176	.006			-.007	-.015	.086			
		(.055)	(.001)	←	←	(.001)	(.007)	(.020)	-316.147	644.295	665.447
		**	***			***	*	***			

Table 13, cont.

Corrected Environmental Stability											
Mean Effect Size	.652 (.064) ***	-	-	-	-	-	-	.215 (.044) ***	-379.335	762.669	769.712
Full Model	-.656 (.270) *	.440 (.045) ***	-.017 (.010)	.023 (.007) ***	.003 (.002)	-.001 (.001)	-.024 (.010) *	.102 (.045) *	-328.563	673.126	701.298
Preferred Model (AIC and BIC)	-.654 (.270) *	.440 (.107) ***	-.018 (.010)	.024 (.006) ***	.002 (.001)	←	-.023 (.010) *	.102 (.045) *	-328.748	671.495	696.145

Note. b_0 represents the intercept; $b_{0\text{-}3}$ represents the coefficient for the first linear segment, $b_{3\text{-}15}$ represents the following segment and so on until $b_{60\text{-}90}$. For simplified models, arrows are used to indicate the parameter estimate that applies during the given developmental period. τ represents the random-effect standard deviation in the between study effect size.
* indicates $p < .05$; ** indicates $p < .01$; *** indicates $p < .001$..

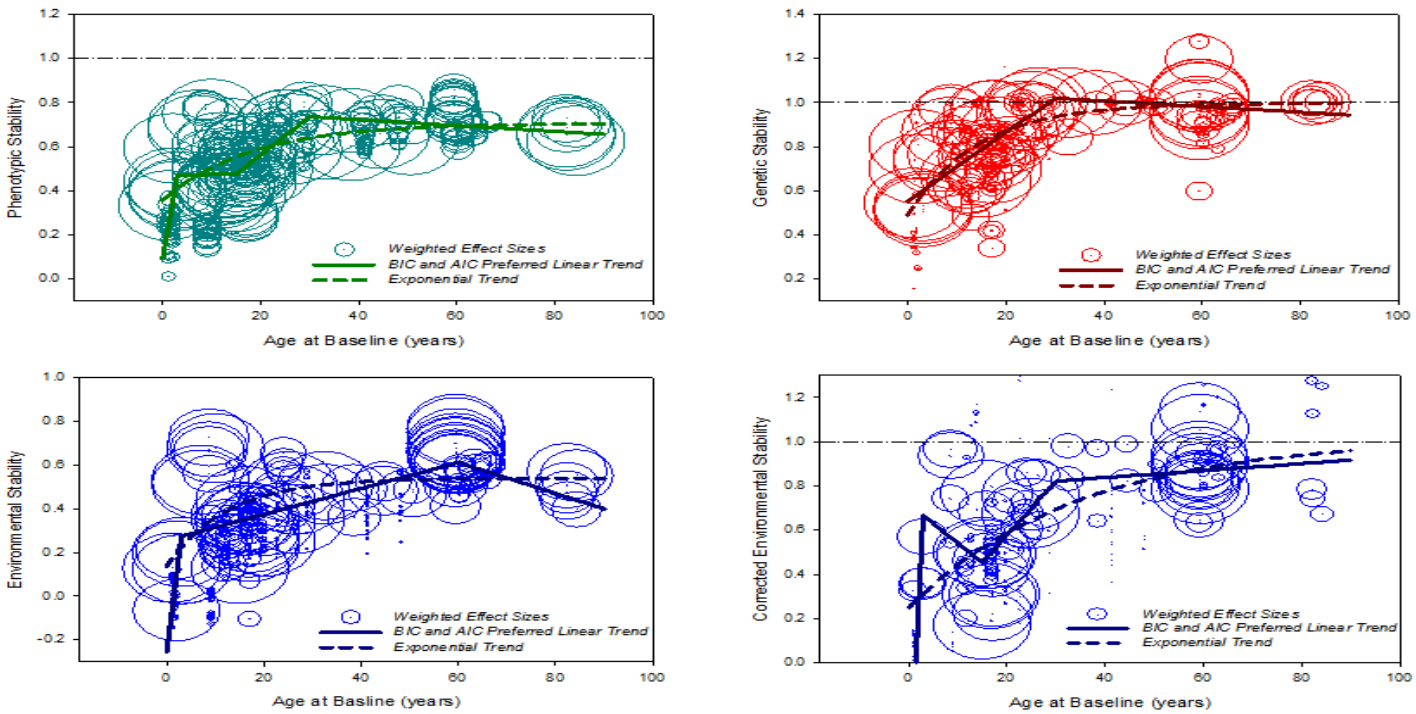

Figure 6. Age-trends in phenotypic, genetic, environmental, and measurement error corrected environmental stability assuming a 5.56 year time lag between assessments. Circles surrounding data points are scaled by the weighting variable (described in Analytic Approach section) such that larger circles 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 behavior 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 to maintain interpretability.

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 lifespan. For phenotypic stability, the full model indicated that in infancy differentially stability was not different from zero (b_0 = .093, p = .47). Stability rose sharply over the first three years of life ($b_{0\text{-}3}$ = .126, p < .05), plateaued during childhood ($b_{3\text{-}15}$ = .001, p = .88), and then resumed increasing substantially until age 30 ($b_{15\text{-}30}$ = .016, p < .001). Following age 30, no slope was significantly different from zero, but there was a small trend towards 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 ($b_{0\text{-}30}$ = .016, p < .001). At this age, genetic stability reached unity and remained nearly perfectly stable across the remainder of the lifespan. The preferred model for environmental stability was slightly more complex. In infancy, environmental stability rises quickly ($b_{0\text{-}3}$ = .176, p < .001), and then continues to rise at a slow, steady rate over the majority of the lifespan ($b_{3\text{-}60}$ = .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 displays 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 or increases slightly ($b_{30\text{-}90}$ = .002, p = .22).

The continuous exponential models are largely consistent with the lifespan trends as indicated by the connected-linear models. Each model indicated growth across the lifespan with an upper asymptote of .71 for phenotypic stability, 1.0 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 stability are .55, .82 and .42, respectively, by age 15. At 30 years of life, the expected phenotypic, genetic and environmental stability 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 environmental stability, on the other hand, continues to rise throughout the lifespan. 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 6. The empirical story is the same across linear, non-linear, 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 14 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 others words, genetic effects exert a constant, moderate effect (b_0 = .358, p < .001) on phenotypic stability across the lifespan. 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 ($b_{0\text{-}30}$ = .008, p < .001) and plateaus afterward ($b_{30\text{-}90}$ = .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 7 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 14. Age trends in the genetic and environmental contribution to phenotypic stability

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

Note. b_0 represents the intercept; $b_{0\text{-}3}$ represents the coefficient for the first linear segment, $b_{3\text{-}15}$ represents the following segment and so on until $b_{60\text{-}90}$. For simplified models, arrows are used to indicate the parameter estimate that applies during the given developmental period. τ represents the random-effect standard deviation in the between study effect size.
* indicates $p < .05$; ** indicates $p < .01$; *** indicates $p < .001$.

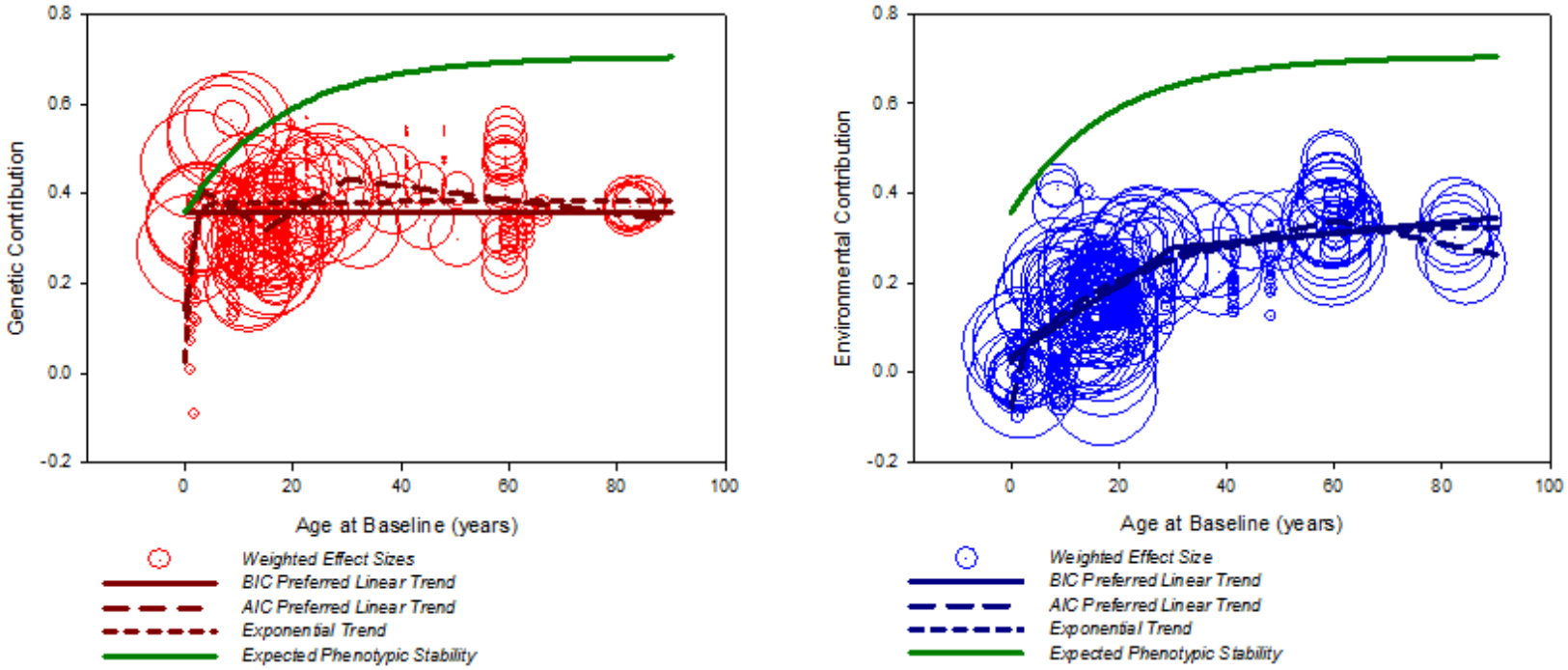

Figure 7. Age-trends in genetic and environmental contributions to phenotypic stability assuming a 5.56 year time lag between assessments. Circles surrounding data points are scaled by the weighting variable (described in Analytic Approach section) such that larger circles carried *more* weight in the analysis.

Table 12 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. 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 lifespan. At age 15, the expected environmental contribution is only .17 correlation units, and it continues to .25 by age 30 and continues to rise past age 60 (expected environmental contribution = .31).

The identified linear and non-linear trends are apparent in Figure 7 with reference to the expected exponential lifespan trend for phenotypic stability plotted in green. In infancy, phenotypic stability equals the genetic contribution to stability, and genes exert an impressive and almost constant influence on stability across the entire lifespan. 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 lifespan. This results in a nearly constant genetic contribution to phenotypic stability. Both environmentality and environmental stability increase across the lifespan. Thus, the combination of environmental effects persisting 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 lifespan.

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 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 12) rather than deviations from a reference category (e.g., the trend for extraversion).

Moderation by Big Five Traits.

Table 15 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, but these deviations amount to about a difference of 1% of variance. Conscientiousness displayed the most consistent deviations from the average trend, and this trait tends to be more environmental, and more stable phenotypically and environmentally (but not genetically), and genes and the environment both contribute more to its stability (as would be expected since it is overall more phenotypically stable). Extraversion tends to be influenced more

by genes, is more stable phenotypically and genetically, and genes contribute more to stability. Agreeableness tends to be more environmental, less stable phenotypically and environmentally, and genes contribute less to stability. No significant differences were found for neuroticism or openness. Again, despite several statistically significant differences between the Big Five traits, the major conclusion is that there are strikingly no pragmatic differences in any of the lifespan trends.

Table 15. Moderation of continuous exponential model by the Big Five

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

Notes. E stands for extraversion. A stands for agreeableness. C stands for conscientiousness. N stands for neuroticism. O stands for openness. The Big Five variables were effects coded (see Methods section), and therefore the parameter estimates represent deviations from the average trend.

* indicates $p < .05$; ** indicates $p < .01$; *** indicates $p < .001$.

Moderation by Self- vs. Other-Report.

Table 16 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.

Table 16. Moderation of continuous exponential model by report format and trait bandwidth

Outcome	Self (.5) vs. Informant (-.5) Report	Broad (.5) vs. Narrow (-.5) Measurement
Heritability	-.007	-.008
	(.001)	(.002)
	***	***
Environmentality	.007	.007
	(.001)	(.002)
	***	***
Corrected Environmentality	.008	.008
	(.001)	(.001)
	***	***
Phenotypic Stability	-.006	-.005
	(.004)	(.002)
		*
Genetic Stability	-.001	-.001
	(.002)	(.002)
Environmental Stability	-.020	-.014
	(.002)	(.007)
	***	
Corrected Environmental Stability	-.033	.018
	(.015)	(.008)
	*	*
Genetic Contribution to Stability	-.011	-.005
	(.002)	(.004)
	***	
Environmental Contribution to Stability	-.003	-.003
	(.000)	(.001)
	***	**

Notes. * indicates $p < .05$; ** indicates $p < .01$; *** indicates $p < .001$.

Moderation by broad vs. narrow measurement

Table 16 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 17 and illustrated in Figures 8-10. Note that

age 9 years is the youngest self-report effect size, and we do not extrapolate the expected trend line to younger ages.

Table 17. Results of continuous exponential models for each developmental outcome

Outcome	b_0	b_1	b_2	$b_{\text{time lag}}$	LL	AIC	BIC
Heritability	.399	-.108	-.043				
	(.073)	(.159)	(.145)	-	-865.505	1737.011	1747.611
	***						
Environmentality	.610	.121	-.033				
	(.082)	(.090)	(.086)	-	-875.896	1757.792	1768.393
	***						
Corrected	.369	2.131	-.259				
Environmentality	(.020)	(2.862)	(.126)	-	-884.588	1775.176	1785.777
	***		*				
Phenotypic	.709	1.283	-.120	-.005			
Stability	(.028)	(.431)	(.027)	(.004)	-1640.088	3288.177	3300.859
	***	**	***				
Genetic Stability	.996	1.185	-.111	-.005			
	(.008)	(.279)	(.018)	(.004)	-1380.600	2769.200	2781.882
	***	***	***				
Environmental	.593	1.104	-.084	-.008			
Stability	(.051)	(.355)	(.028)	(.004)	-1412.135	2832.269	2844.951
	***	**	**	*			
Corrected	.901	1.598	-.079	-.018			
Environmental	(.017)	(.357)	(.013)	(.007)	-1328.502	2665.003	2677.685
Stability	***	***	***	*			
Genetic	.395	.393	-.111	.006			
Contribution to	(.016)	(.145)	(.033)	(.002)	-1405.376	2818.751	2831.433
Stability	***	**	**	*			
Environmental	.325	.647	-.083	-.007			
Contribution to	(.020)	(.178)	(.026)	(.004)	-1505.100	3018.199	3030.881
Stability	***	***	**	*			

Note. * indicates $p < .05$; ** indicates $p < .01$; *** indicates $p < .001$.

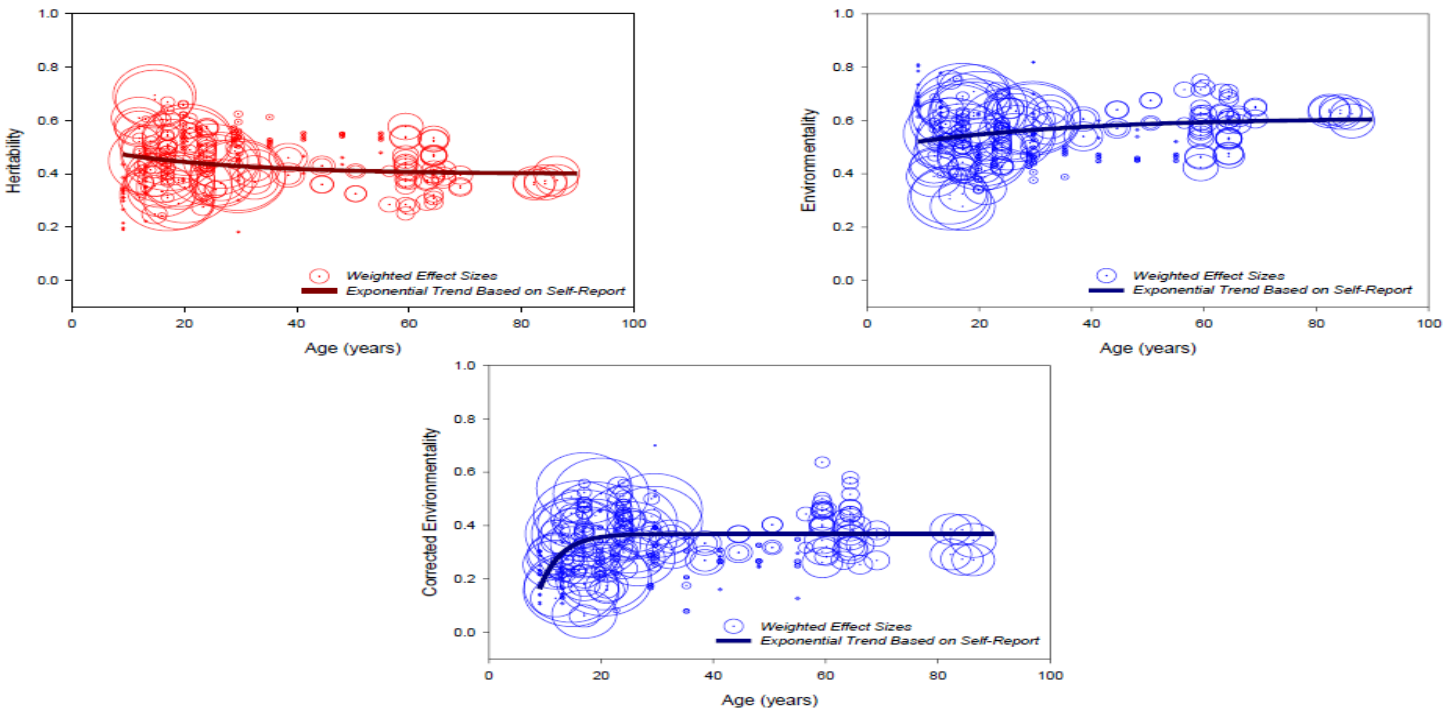

Figure 8. Age-trends in heritability, environmentality, and measurement error corrected environmentality based *only* on self-report data. Circles surrounding data points are scaled by the weighting variable (described in Analytic Approach section) such that larger circles carried more weight in the analysis.

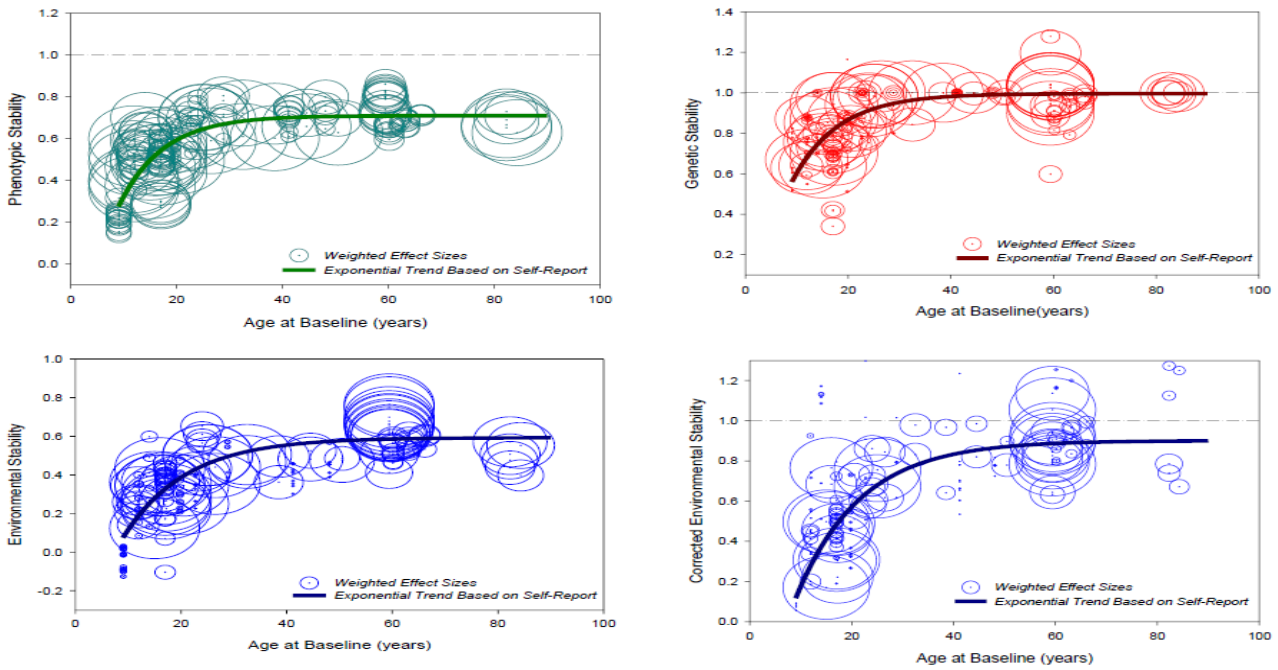

Figure 9. Age-trends in phenotypic, genetic, environmental, and measurement error corrected environmental stability based *only* on self-report data assuming a 5.56 year time lag between assessments. Circles surrounding data points are scaled by the weighting variable (described in Analytic Approach section) 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 behavior 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 to maintain interpretability.

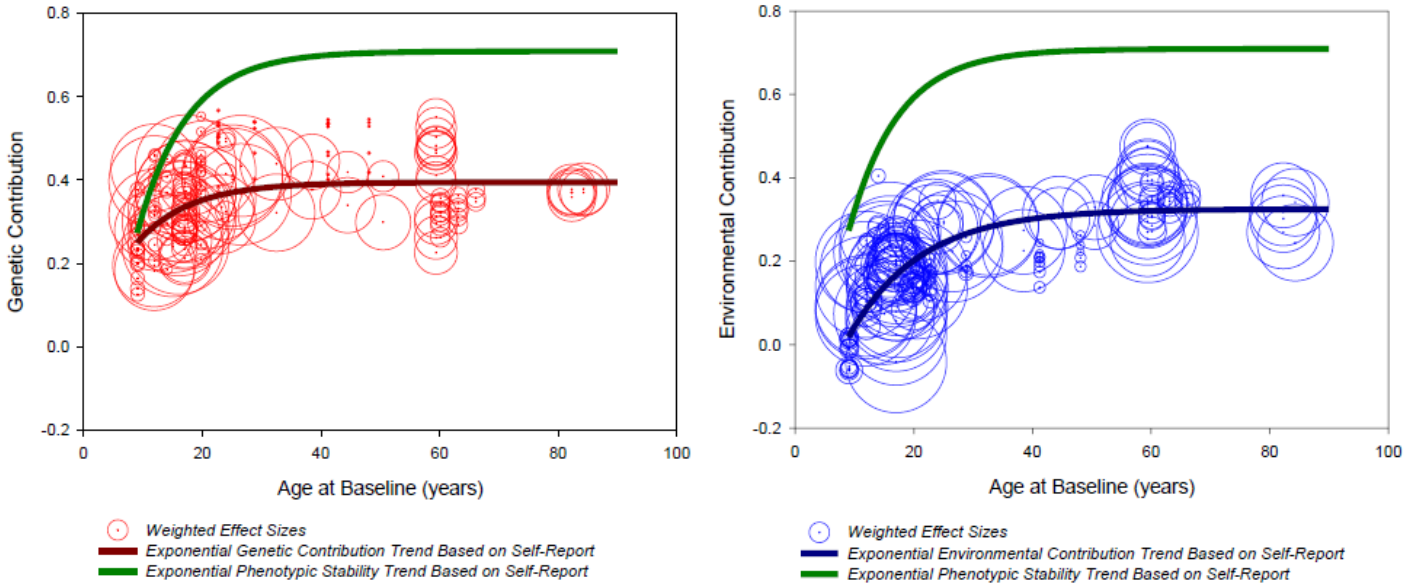

Figure 10. Age-trends in genetic and environmental contributions to phenotypic stability assuming a 5.56 year time lag between assessments. Circles surrounding data points are scaled by the weighting variable (described in Analytic Approach section) such that larger circles carried *more* weight in the analysis.

Consistent with the results of analyses of the complete dataset, results of analyses of only self-report data indicate that phenotypic, genetic, and environmental stabilities increase significantly with age, particularly up until about age 30 years (Figure 9). 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 dataset, indicating that they are not artifacts of blending information from self- and informant- report formats. Age-related trends for heritability and environmentality derived from analyses of only self-only report data (Figure 8) are weaker than for those derived from analyses of the complete dataset. For analyses of only self-report data, heritability does appear to decline (and environmentality increase) slightly early in life, but this trend is not statistically significant, and the predominant pattern is that of consistency of genetic and environmental influences on personality across the lifespan. Corrected environmentality did significantly increase with age, but this increase was modest and plateaued quickly. Of course, if the most dramatic changes in heritability and environmentality 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 dataset, analyses of only self-report data indicate that age-related increases in phenotypic stability are predominantly attributable to increases in environmental contributions (Figure 10). 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 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 effect from the continuous exponential models. In instances in which the standard errors were positively skewed (this 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, environmentality, and corrected

environmentality. For heritability, the distribution of residual effect sizes was negatively skewed with a few highly negative values (for environmentality, the distribution of effect sizes was positively skewed). These effect sizes were primarily drawn from the first years of life. Returning to Figure 5, it is apparent that there are several effect sizes that indicate low heritability during this period of the lifespan. 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 datapoints 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, environmentality, corrected environmentality, 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 dataset. The lifespan trends from the resulting model were essentially the same as those reported in Table 12. Therefore, while there may be some publication bias for low precision studies, the meta-analytic estimates reported in this manuscript 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 lifespan. 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 others (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 genes 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, genes and environments 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 lifespan. 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 lifespan. 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 three 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 lifespan 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 lifespan. 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 et al., 1990), heritability of personality decreased with age, and environmental influences increased. Moreover, environmental influences increased in stability across development, evincing impressive levels in adulthood. Further, 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-by-environment interaction and gene-environment correlation are largely consistent with the current results. In early childhood, it may be the case that substantial and extensive gene-by-environment interaction takes place based on the shared environment provided by parents. This would explain the very large heritability of personality in early childhood. As children mature and begin to leave this shared environment, heritability begins to decrease indicating that the gene-by-environment interactions may not be particularly lasting (i.e., a scarring effect that would persist beyond the environmental exposure), but rather frequently recur over early development. This is likely to be the case as aspects of the objectively shared environment tend to be stable during the child's early life (e.g., the parenting they receive or their macro-environmental resources). However, this process would entail a decrease

in genetic stability as children moved away from environments that they differentially responded to on the basis of genotype. We found the opposite trend, a fairly constant increase in genetic stability through childhood. Gene-environment correlation may explain the increase in genetic stability, as children gain the autonomy to increasingly 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 be "turned on," and the resulting phenotype becomes less influenced by novel gene expression over time. Because genetic effects decrease in importance but also increase in stability, it is possible that any or all of these mechanisms occur simultaneously.

Gene-by-nonshared environment interaction may explain the lifespan 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 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 means that heritability would decrease with age and 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 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-by-environment interactions "are ubiquitous in nature" and evolutionarily adaptive (Penke, Denissen, & Miller, 2007, p. 578). Phenotypic plasticity in response to environmental circumstances and resources allow organisms to adjust to environmental variation (Dall, Houston, & McNamara, 2004). This may explain the extremely high heritability of personality in very early childhood as 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 answer 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 behavior 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 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. Costa and McCrae (2008) have similarly commented that "ambitious attempts to pin down substantive contributions of the non-shared environment have largely failed" (p. 168). Importantly, our findings indicate a substantial proportion of variance in personality traits is influenced by environmental experiences that act 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 should be considered an encouraging prospect for future investigation of measurable environments. Longitudinal, genetically informative samples will likely be necessary to pin point 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 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 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 ability, non-additive 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 practices. Non-additive 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 accumulate, 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 lifespan 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 appears 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 macro-environmental 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-by-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.

Strengths and Limitations

This study is the first to empirically demonstrate the genetic and environmental stability trends of personality across the entire lifespan. By aggregating data and applying meta-analytic techniques, we have provided a provocative picture of what personality

development looks like. 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 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, this allows for a direct estimation of

the underlying data. When only reduced or trimmed models were reported, the group correlations are 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 lifespan. This 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 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. However, it is also possible that parents tend to exaggerate differences between their (non-monozygotic) 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, this 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 α as an estimate of internal consistency in order to correct the estimates of environmentality and environmental stability for measurement error. However, Cronbach's α is an overestimate of measurement error in several situations, particularly when the measure is multidimensional (Schmidt, 1996). Thus, our models may have overcorrected for measurement error. With this 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 non-linear manner as more time passes between assessments towards 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 lifespan trends, and the quadratic term tended to be small in magnitude[7]. While the current article was specifically focused on age-based lifespan trends in stability, it will be valuable for future research to focus specifically on time-based trends in stability, in order to determine the most appropriate functional form of

[7] 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.

stability decay, and whether such decay interactions 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 macro-environmental context 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 lifespan trend (Bell, 1953). The alternative to this limiting assumption is to track a single cohort through the lifespan to examine the differential stability of personality, but no such sample is available.

Conclusion

Individual differences in patterns of thoughts, feelings, and behaviors 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 lifespan 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 lifespan and fully explains 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 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 such gene-environment interplay carries forward in time.

Chapter 3: Child Characteristics and Parental Educational Expectations – Evidence for Transmission with Transaction

Educational attainment predicts key life outcomes, such as income (Day & Newburger, 2002) and health (Montez, Hummer, Hayward, Woo, & Rogers, 2011). *Educational expectations* (i.e., expecting to continue on an educational track, rather than expecting to pursue other vocational options) are associated with actual educational attainment and academic achievement, even after controlling for a number of family and individual confounds (Alexander, Entwisle, & Bedinger, 1994). This result has been found in a number of high-quality, longitudinal, nationally representative datasets (Jacob & Linkow, 2011), in high risk samples (Ou & Reynolds, 2008), and even as a mediator of program effects in a randomized controlled experiment (Purtell & McLoyd, 2013). Having reviewed much of this literature, Schneider and Stevenson (1999) concluded, "One of the most important early predictors of social mobility is how much schooling an adolescent expects to obtain" (p. 4). Children are thought to form such educational expectations largely in response to parental inputs (Jacobs & Eccles, 2000; Schneider, Keesler, & Morlock, 2010). However, much work in child development more generally (e.g., Bell, 1968) has highlighted the importance of reciprocal or *transactional* processes between children and their environments. The current project demonstrates that both parent-to-child and child-to-parent effects influence academic development and expectations, that these effects occur even before children enter school, and that the transactional process is sensitive to child motivation and problem behavior.

THEORIES OF EXPECTATIONS AND ACADEMIC ATTAINMENT

Sociologists and psychologists have examined the relation between educational expectations and educational attainment in parallel literatures. Both of these literatures emphasize a *transmission* model, in which parents act as active broadcasters of academic beliefs and children act as passive receivers. Building on the sociological work of Blau and Duncan (1967), the *status attainment model* (Sewell & Hauser, 1972; 1980) notes that society is stratified in terms of background characteristics, such as race or socioeconomic status, which in turn reproduce status inequalities in successive generations. However, there are a number of intervening mechanisms between socioeconomic background and academic and occupational success. For example, Sewell and Hauser (1972) hypothesized that the influence of significant others (parents and peers) and academic expectations partially mediates the influence that family background characteristics exert on attained status. Rather than society selecting individuals into various status levels based solely on ascribed factors, individuals can obtain social mobility through social psychological mechanisms (Sewell & Hauser, 1980). For instance, optimistic parental educational expectations may help a child achieve greater academic success than would be predicted simply based on his or her families' socioeconomic background, whereas pessimistic parental educational expectations may influence a child to achieve to a lesser extent. Under this perspective, child academic trajectories and interactions with the educational system reflect the influence of internalized parental beliefs

The expectancy-value (E-V) model is one of the primary psychological frameworks for understanding the interrelations among psychological characteristics leading to motivation and task success (Eccles & Wigfield, 2002; Nagengast, Marsh, Scalas, Xu, Han, Trautwein, 2011). Under this framework, the primary determinants of motivation to complete a task are the expectation that the task can be completed and the value of completing the task (Jacobs & Eccles, 2000, see also Bandura, 1986; Zimmerman, Bandura, & Martinez-Ponz, 1992). In the realm of academic motivation, perceived academic competence and belief in the worth of school represent expectancies and values, respectively. Parents are thought to instill in their children perceptions about the value of schooling and about the children's ability to succeed in school based, in part, on the parents' own expectations regarding the level of credentials their children will obtain. In turn, parental expectations are thought to be sensitive to child development. Parents of children that show promise academically may raise their expectations in response to this new information and provide academic stimulation in the form of increased involvement in the child's school life. Thus, the E-V model implies a highly complex, dynamic, and reciprocal relationship between parents and children. However, as we describe below, this level of theoretical nuance, particularly the role of child characteristics and behaviors in shaping parental expectations, is infrequently incorporated into the empirical literature.

TRANSMISSION OF ACADEMIC BELIEFS

A key component of the empirical work on educational expectations is the assumption of a causal effect of *parental* expectations on children's attainment. Under this assumption, parents instill levels of educational expectations in their children, which are then internalized to inform academic self-concepts. Consistent with this assumption, Gonzalez-Pienda, Nunez, Gonzalez-Pumariega, Alvarez, Roces, & Garcia (2002) found that parental expectations were significantly associated with child beliefs regarding competence and academic aptitude. These latter two variables were significantly associated with achievement and mediated the effect of parental expectations. That the associations between parental educational expectations and child outcomes are statistically mediated through child expectations, task value and academic self-concept has been well replicated (Beal & Crockett, 2010; Bleeker & Jacobs, 2004; Fredricks & Eccles, 2002; Frome & Eccles, 1998; Neuenschwander, Vida, Garrett, & Eccles, 2007; Simpkins, Fredricks, and Eccles, 2012). Validating the utility of this line of research, Harackiewicz, Rozek, Hulleman, and Hyde (2012) developed a successful intervention based on the E-V model that instructed parents on effective ways to show their children the value of science related courses for their life goals. The high-school children of parents in the experimental group completed significantly more science coursework than the children in the control group, signifying the importance of parents as academic motivators.

Modeling parents as an exogenous influence on child development, as is common in expectations research, makes intuitive sense when one considers that many of the other influences on academic success are ascribed factors (e.g., socioeconomic status).

Supporting this position, Andrew and Hauser (2011) found that adolescent students largely adopt levels of educational expectations based on social background characteristics and adapt to academic feedback (i.e., grades) very modestly. Similarly, Tynkkyen, Tolvanen, and Salmela-Aro (2012) tracked adolescent's trajectories of educational expectations over five years as a function of social background and parental expectations. Parental expectations were significantly associated with trajectories of child academic expectations, and importantly, developing along different expectation trajectories resulted in disparities in achievement for the children. In these studies parental expectations are assumed to take both chronological and causal precedence over child variables, and the dynamic nature of the parent-child relationship and development of expectations may be even more obscured due to the relatively old age of students.

The results of the empirical studies reviewed above have led researchers to draw strong conclusions concerning the transmissive properties of the influence that parents have on their children. For example, the assertion that "educational expectations that parents have for their children represent one of the key mechanisms through which parents influence their children's schooling careers" implies an underlying transmission process (Schneider et al., 2010, p. 253). Jacobs and Eccles (2000) claim that "the direction of influence for perceptions of competence is from parents to children" (p. 420). Following in this tradition, Simpkins et al. (2012) justified their conceptual model with parenting beliefs and behaviors preceding and independent of child beliefs and behaviors based on past research indicating that "mothers' beliefs shape child development" (p. 1020). Conceptual or path diagrams with similar causal ordering claims are found in

several expectations studies (Bleeker & Jacobs, 2004; Frome & Eccles, 1998; Gonzalez-Pienda et al, 2002; Neuenschwander et al., 2007). While the extant evidence certainly is consistent with parent-to-child effects, an outstanding and conspicuously neglected empirical question is whether parental educational expectations are subject to dynamic and reciprocal feedback from children. Importantly, while such dynamics have been regularly incorporated into *theoretical* publications concerning the E-V model (see Jacobs & Eccles, 2000, p. 423), they have, to date, seldom been explicitly examined in empirical designs.

TRANSACTIONAL PROCESSES BETWEEN PARENTS AND CHILDREN

Transaction, as opposed to transmission, represents an elaborated framework for understanding socialization (e.g., Sameroff, 2009, but see also Bronfenbrenner & Ceci, 1994; Collins, Maccoby, Steinberg, Hetherington & Bornstein, 2000; Lerner & Busch-Rossnagel, 1981; Scarr & McCartney, 1983; Tucker-Drob, Briley, & Harden, 2013). While transmission models view parents as broadcasters and children as receivers, transactional models emphasize the dynamic roles found in the socialization process. Bell (1968) was one of the earliest researchers to argue that children, even infants, play an active role in influencing the parenting that they receive, and thereby their own development. Transaction implies that simple associations between a parenting behavior and a child outcome are causally ambiguous, because it is unknown whether the parent influenced the child or the child influenced the parent. Causal ambiguity in studies of parenting also occurs because parents pass on to their children both an environment and

genetic predispositions. The research methods used in empirical studies must be able to address these types of alternative hypotheses.

Work in developmental behavior genetics offers an empirically tractable solution to this problem by allowing for the possibility that children's genetically influenced behaviors and dispositions influence the types and quality of experiences that they evoke from others (Plomin, DeFries, & Loehlin, 1977). This process is termed gene-environment correlation to refer to the correlation that arises between children's genotypes and the environments that they receive.

Gene-environment correlation provides several avenues for the relationship between parental educational expectations and child academic beliefs to occur in addition to unidirectional transmission of values. Child characteristics and behaviors, such as motivations, abilities and self-concepts, may be subject to genetic influences. Parents may be sensitive to these genetically-influenced characteristics and adjust their expectations accordingly. This is one potential mechanism whereby the genetic predispositions of the child are able to get "out of the skin" and influence the environment. Thus, evidence that parental educational expectations are "heritable" on the part of children would indicate child-to-parent effects. Two recent meta-analyses found that child genetic effects account for roughly 20% of the variance in parenting behaviors depending on the variable (Avinun & Knafo, 2013; Kendler & Baker, 2007). McAdams, Gregory, and Eley (2013) attempted to explain why parenting variables are "heritable" by examining the correlation between genes influencing variation in child attributes and parenting behavior. They found that adolescent genetic influences on a set of maladaptive

traits (e.g., oppositionality and depression) were able to explain the genetic influences on parenting. However, the extent to which this pattern of gene-environment correlation applies to variation in parental educational expectations and their associations with child achievement is unknown.

There are a handful of studies that provide preliminary support for reciprocal parent-child transactions involving educational expectations or academic beliefs. Zhang, Haddad, Torres, and Chen (2011) used cross-lagged path models to simultaneously control for parent and child characteristics. This has the desirable effect of modeling the prospective influence of student expectations, parent expectations, and academic achievement on one another over time, above and beyond baseline levels of each outcome. Consistent with a transactional hypothesis, significant bidirectional cross-lagged paths were found between each variable. However, this study tracked the developmental process relatively late in the academic careers of the participants (between 8^{th} and 12^{th} grades) rendering it unclear how early this process begins. Wang (2012) detected similar reciprocal relations between student beliefs, grades, and the classroom environment. Marsh and colleagues (Marsh & O'Mara, 2008; Marsh & Martin, 2011; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005) demonstrated reciprocal effects between academic self-concept, interests, grades, and academic achievement. However, they did not examine parental educational expectations, or any other parent behaviors for that matter.

PROPOSED TRANSACTIONAL MODEL UNDER INVESTIGATION

For the current study, we track the transactional relations between three classes of variables: child academic behavior, child academic achievement outcomes, and parental educational expectations. By child academic behavior, we mean behavioral tendencies relevant to academic success or difficulty. For example, a child who diligently completes chores or pays attention may inspire higher expectations. Conversely, a child who creates trouble or has emotional outbursts may lower expectations. By academic achievement outcomes, we mean performance on standardized tests of math and reading. Objective test scores are one of the strongest correlates of educational attainment (Strenze, 2007). Similar to child academic behavior, a particularly bright child may enlist increasing parental expectations. Parental educational expectations are predicted to positively influence beneficial academic behaviors and cognitive development and hinder problematic academic behaviors. Additionally, greater academic achievement is likely to reinforce positive academic behaviors and reduce problematic behaviors. Child positive behaviors likely enable successful learning, and problem behaviors likely interfere with academic achievement. Therefore, we posit a fully interactive transactional model where child academic behaviors, child cognitive development and parental educational expectations each influence the other constructs over development. Furthermore, as nearly all reliably measured psychological variables are subject to genetic influences (Turkheimer, 2000), we expect transactional processes to result in children's

environmental circumstances (i.e., parental educational expectations) becoming tied to their genotypes.

A number of processes may mediate or moderate this transactional model. First, parental educational expectations are a distal factor and likely influence the more proximal interactions between parent and child, such as involvement or cognitive stimulation (Wang & Sheikh-Khalil, 2013). We will examine whether these types of proximal behaviors can mediate a portion of the influence of parental expectations on child outcomes. Second, the role of educational expectations may differ as a function of sociodemographic factors, such as race (Hanson, 1994; Mickelson, 1990), gender (DiPrete & Jennings, 2012; Jacobs & Eccles, 1992; McWhirter, 1997), and socioeconomic status (Davis-Kean, 2005). We will evaluate whether the transactional processes differ across sociodemographic dimensions.

GOALS OF THE PRESENT STUDY

Our study has two primary goals. First, we use behavioral genetic models of data from twins to evaluate whether parental educational expectations are associated with genetic differences in their children. The current study follows the recommendations of a number of researchers to incorporate genetically informed methods into the study of the social environment (D'Onofrio, Lahey, Turkheimer, & Lichtensten, 2013; Harden, 2013; Reiss, 2003). As recommended by these authors, we do so to unravel complex family dynamics, rather than to simply establish the heritabilities of the outcomes. Although recent reviews (e.g., Crosnoe & Johnson, 2011; Schneider et al., 2010) of the relevant

developmental literature have noted the importance of integrating genetic thinking with socialization models, we are aware of no study that has used a genetically informative sample to evaluate associations between expectations and child academic behaviors or achievement. Second, we use cross-lagged longitudinal models to evaluate specific transactional processes that occur between child academic behavior, child academic achievement outcomes and parental educational expectations. Importantly, we test whether the transactional processes of interest can be detected even prior to kindergarten entry. The majority of research on educational expectations has focused on children relatively late in their academic careers, but there is evidence of transactional processes that influence cognitive development before children even enter the educational system (Lugo-Gil & LeMonda, 2008; Tucker-Drob & Harden, 2012a). It is possible that investigators focusing on the middle school and high school years may be searching in the wrong place for the origins of social stratification in academic achievement and educational attainment (Barnett, 1995; Downey, von Hippel, & Broh, 2004; Duncan et al., 2007; Heckman, 2006; McLoyd, 1998; Tucker-Drob, Rhemtulla, Harden, Turkheimer, & Fask, 2011).

METHOD

Sample

Data were drawn from the Early Childhood Longitudinal Study – Birth Cohort (ECLS-B) and Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K). These separate datasets are ideal for analyzing the questions posed in that they contain

high quality assessments of children's development and their environments (Snow et al., 2009; Tourangeau, Nord, Lê, Sorongon, & Najarian, 2009). ECLS-B is nationally representative of United States children born in 2001, and ECLS-K is nationally representative of the kindergarten cohort of 1998. The data used in the current study was the age 4 (collected in 2005 and 2006) and kindergarten waves (collected in 2006) of ECLS-B. The data from ECLS-K included the fall kindergarten (collected in 1998), spring first grade (collected in 2000), third grade (collected in 2002), and fifth grade (collected in 2004) waves. Waves not listed above were omitted due to limited measure or data availability. The initial wave of ECLS-B recruited 10,650[8] parents to participate, and ECLS-K recruited 22,666 children. The racial composition of the ECLS-B sample was 41% White, 16% African American, 21% Hispanic, and 11% Asian. The racial composition of ECLS-K was 51% White, 14% African America, 16% Hispanic, and 6% Asian. The remaining participants were identified as Pacific Islander, Native American, multiracial or unknown. Males represented 51% of both the ECLS-B and ECLS-K samples. Our behavioral genetic models were fit to the twin subsample of ECLS-B. Data were available for 1,200 twins. The racial composition of the subsample was 61% White, 16% African American, 16% Hispanic, and 3% Asian with an equal percentage of males and females.

[8] ECLS-B confidentiality requirements state that all reported sample sizes must be rounded to the nearest 50.

Measures

Zygosity

Twin zygosity was ascertained by trained coders at the second wave of the ECLS-B data collection. Twins were rated on the similarity of their physical appearance (e.g., hair texture, eye color, ear lobe shape). These items ranged from 1 (no difference) to 3 (clear difference). Using the procedure described in Tucker-Drob et al. (2011), we computed sum scores from the six items, which ranged from 6 to 18. Twin pairs with zygosity scores below 8 were classified as monozygotic. Same-sex twin pairs classified as dizygotic were removed from the sample if the parents reported a medical reason for the twin's dissimilarity. Previous research has found that zygosity diagnoses obtained from such physical similarity rating approaches are over 90% accurate when validated with biospecimens (Forget-Dubois et al., 2003). We excluded any twin pair in which either twin had been diagnosed with a developmental delay, mobility disorder, or autism (13% of original pairs). Additionally, we excluded any twin pair discordant for pre-school care arrangement as this was a very small proportion of twin pairs (<1% of original pairs). The results were very similar when these excluded pairs were included. Our final sample was composed of 29% monozygotic twin pairs, 35% same-sex dizygotic twin pairs and 36% opposite-sex dizygotic twin pairs.

Parental Educational Expectations

At each wave, parents were asked what degree they expected their children to achieve. The response options were to receive less than a high school diploma, to graduate from high school, to attend two or more years of college, to finish a 4-or-5 year college degree, to earn a master's degree or equivalent, and to get a Ph.D., MD, or other higher degree. The survey items were equivalent across ECLS-K and ECLS-B.

Academic Behavior

Academic behavior, operationalized in terms of approaches towards learning and problem behavior, was collected at each measurement wave. Approaches towards learning items were chosen to represent active, child centered behaviors that would facilitate transactions with the academic environment, and problem behaviors were chosen as potential sources of interference for transactions. In ECLS-B, parents reported on the extent to which their children are eager to learn, pay attention, work independently, and work until finished. These items were used to assess approaches towards learning (see Tucker-Drob & Harden, 2012b). Parents additionally reported on the extent to which their children were aggressive, angry, impulsive, overly active, have temper tantrums, annoy other children, and destroy other children's belongings. These items were used to assess problem behaviors (see Tucker-Drob & Harden, 2013). Each item was rated on a scale from 1 (never) to 5 (very often). Average scores were calculated for each individual.

Because parent report of educational expectations is a primary study variable, we were concerned that method variance would be shared with parent report of academic

behavior. To complement parent report of academic behavior, we also included teacher report of the same items. However, teacher report was only obtained for children that attended a formal childcare setting at age 4. This reduces the twin sample size by 33% and represents a serious loss in power. In the ECLS-B dataset, children that attend pre-school are systematically different from those that do not in terms of their academic growth (Tucker-Drob, 2012). Further, parents and teachers may have access to different information about the focal children. In light of these concerns, we evaluate the convergent validity of these sources of information both phenotypically and at the behavioral genetic level. We analyzed both variables independently in order to detect any potential differences in the pattern of association with achievement and parental expectations.

In ECLS-K, teacher reports of approaches towards learning and externalizing behavior were used to represent similar constructs. Scale scores were computed by the ECLS-K research team that ranged from 1 (never) to 4 (very often). Upon initial inspection of the data, we determined that the approaches towards learning scales were negatively skewed and the problem behavior scales were positively skewed. Transformations were conducted that minimized skew. For the ECLS-B variables, this involved taking the square root of each score. The ECLS-K approaches towards learning scores were also transformed by the square root, but the externalizing scores displayed larger skew and taking the inverse minimized skew. Reliability was acceptable in both datasets and for both measures. Reliability of approaches towards learning ranged from .66 to .68 (parent report) and .79 to .80 (teacher report) in ECLS-B and from .89 to .91 in

ECLS-K. Reliability of problem behaviors ranged from .78 to .80 (parent report) and .79 to .83 (teacher report) in ECLS-B and from .86 to .90 in ECLS-K.

Academic Outcomes

Both ECLS databases contain extensively developed math and reading achievement scores collected at each data wave. The test materials were modified over the course of the study to account for the dramatic gains in general ability across development. The subject matter and specific skills required for the tests changed with age. Item response theory models were applied to the raw data to calculate comparable scores for each participant regardless of wave of assessment. For a complete description of the test procedures and application of the scoring procedure, see Snow et al. (2009) and Tourangeau et al. (2009). Estimated reliability of the indicators was high in all datasets and waves of data. In ECLS-B, reliability of math and reading achievement variables ranged from .84 to .92 across waves. In ECLS-K, reliability of math and reading achievement variables ranged from .89 to .96.

Potential Mediators

To assess mediators that may be more proximal behavioral manifestations of educational expectations, we constructed composites in ECLS-K that indicate parental involvement and parental stimulation of cognitive development. Parental involvement was indicated by whether the parent attended an open house, a PTA meeting, a parent-teacher conference, a school event, acted as a school volunteer, participated in

fundraising, and had met the child's teacher. Parental stimulation of cognitive development was indicated by parent report of the frequency that the family told stories, sang songs, did art, played games, taught about nature, built things, played sports, practiced numbers, and read together. These indicators were only available at the first and third grade waves. The variables available in ECLS-B are less extensive, particularly those for school involvement, and we therefore chose to focus the meditation analysis on ELCS-K. Reliability of parental involvement ranged from .59 to .60, and reliability of parental stimulation ranged from .71 to .73.

Sociodemographic Controls and Moderators

We included a number of sociodemographic variables including maternal age, type of pre-school care (none, relative care, nonrelative care, center, or head start), disability status, child gender, child race, and family socioeconomic status. As described below, these variables acted as both control variables and moderators in the longitudinal analyses.

Analytic Approach

We used behavior genetic models that capitalize on the known differences in genetic similarity of monozygotic and dizygotic twins reared together to make inferences about the effect of additive genetic influences (A), shared environmental influences (C) that operate to make twin pairs more similar to one another, and nonshared environmental influences (E) that operate to make twin pairs more dissimilar to one another. The E

estimate also includes measurement error, which (by definition) is not correlated across twins, and thus renders twins dissimilar. Figure 11.A. illustrated this approach as a structural equation model that is estimated as a multiple group model for monozygotic and dizygotic twins. An outcome is measured independently for each twin (indicated by the squares labeled Twin1 and Twin2), and the variance in this outcome is fully decomposed into the effects of A, C, and E. The latent factors A1 and A2 represent the specific genotypes of the twins. For monozygotic twins, the correlation between these genotypes is fixed to 1.0, reflecting the assumption that monozygotic twins share nearly identical genotypes. For dizygotic twins, this correlation is fixed to 0.5, reflecting the assumption that dizygotic twins share, on average, 50% of segregating genes. Because the twins share a common rearing environment, the latent variable C is represented by a single factor across both twins. Finally, the E factor, representing unique environmental effects on the outcome, is not correlated across twins. The parameters *a*, *c*, and *e* are constrained to be equal across the twin pair. As the portion of the model representing Twin2 is largely redundant, we will only display the model relevant to Twin1 (pathways indicated by solid rather than dashed lines) in subsequent path diagrams.

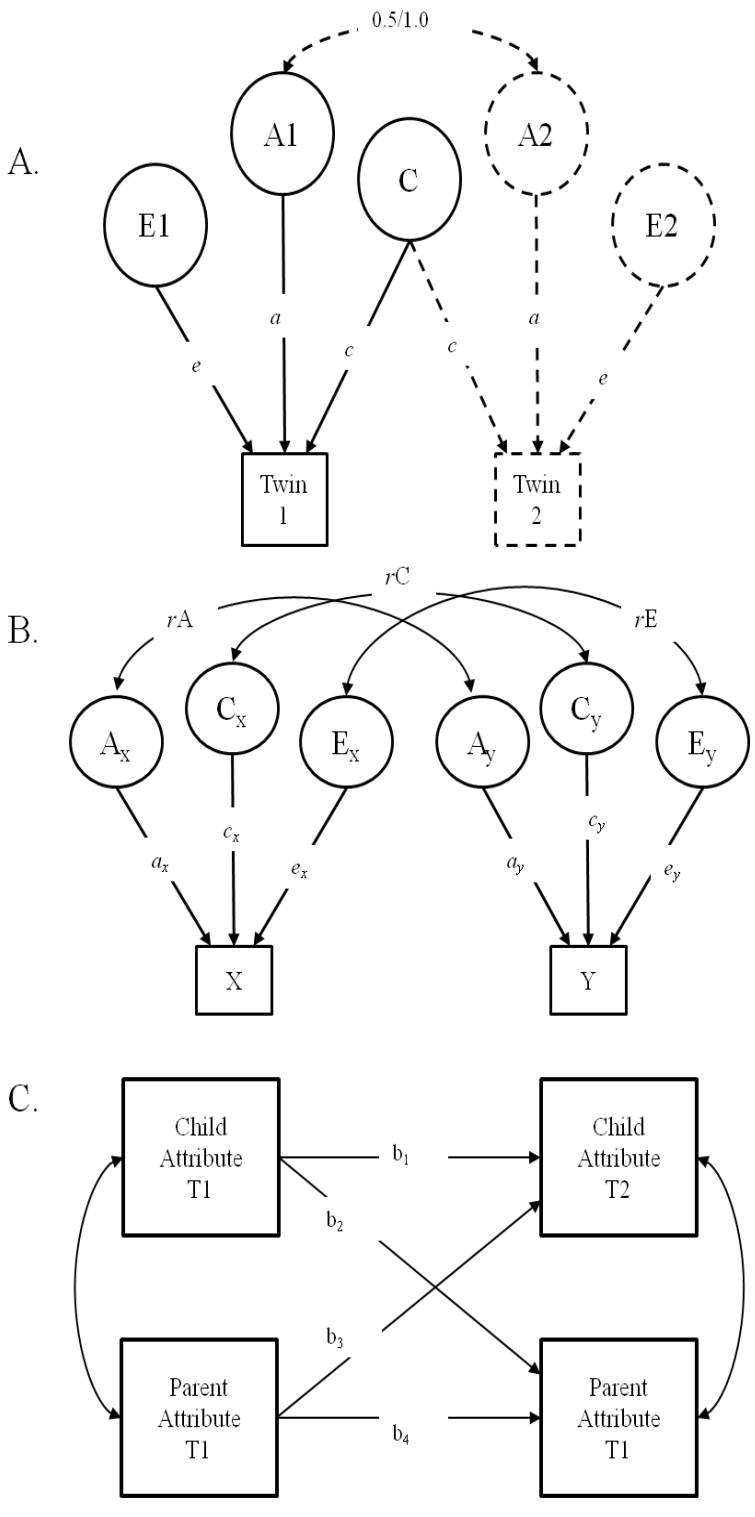

Figure 11. Example model types used in the current study represented as structural equation models with manifest (squares) and latent (circles) variables. **A.** Univariate behavior genetic model for monozygotic and dizygotic twins reared together. In this model, the variance in a measured outcome is decomposed into that which is due to latent additive genetic (A), shared environmental (C), and nonshared environmental (E) factors. Parameters are constrained to be equal across twins. The model is a multiple group model with the correlation between twin1's genetic factor and twin2's genetic factor set to 1.0 for monozygotic twins and 0.5 for dizygotic twins. Portions of the model represented with dashed lines are largely redundant and not included in later models. **B.** Correlated factors model. In this model, the correlation between two outcomes is decomposed into genetic (*r*A), shared environmental (*r*C), and nonshared environmental (*r*E) correlations. This assesses the extent to which the same or different genetic or environmental factors influence the outcomes. **C.** Cross-lagged path model. In this model, the longitudinal stability of the outcomes is indicated by the autoregressive pathways (b_1 and b_4). The cross-paths (b_2 and b_3) establish the directionality of effects between child and parent.

When the variable under investigation is a measure of the environment, the *a* parameter represents the extent to which individuals select or evoke environmental experiences congruent with their genotype (i.e., active or evocative gene-environment correlation; see Avinun & Knafo, 2013). If this parameter is significant, it would suggest that parents are sensitive to child-to-parent influences. The *c* parameter represents the extent to which environmental influences are experienced equally by both members of the twin pair regardless of zygosity, because of equal parental treatment or common family background. Finally, the *e* parameter represents unequal within-family treatment that is not associated with genotypic differences of the children.

The latent C factor includes all unmeasured influences that operate to make twins living in the same family more similar to one another regardless of zygosity. A number of measured family-level variables are included in the ECLS-B dataset. Incorporating these variables in the behavioral genetic model has the effect of lowering the estimate of C by a proportion equivalent to the amount of variance accounted for by the measured variable (Turkheimer, D'Onofrio, Maes, & Eaves, 2005). If parental educational expectations, child academic achievement, or child academic behavior have a substantial shared environmental influence, we will attempt to explain this influence with known predictors of academically relevant variables including socioeconomic status, race, pre-school type, and maternal age (Bradley & Corwyn, 2002; Tucker-Drob, 2012). Importantly, these variables are *necessarily* shared environmental variables because of the way they are measured; that is they are the same for both members of a twin pair.

Multivariate behavior genetic models draw information from cross-twin cross-variable correlations to index the extent to which genetic and environmental influences on educational expectations, academic behaviors and academic outcomes are shared. These are represented as correlations between latent genetic (*r*A), shared environmental (*r*C) and nonshared environmental (*r*E) factors in the *correlated factors* model depicted in Figure 11.B. Importantly, this figure illustrates only one member of the twin pair, and the latent variables reflect the genetic and environmental influences on variables X and Y rather than Twin1 and Twin2.

The model presented in Figure 11.B. is essentially the behavior genetic extension of a correlation coefficient. That is to say, the entire phenotypic correlation between two variables has been decomposed into genetic and environmental components. To determine the amount that genetic effects mediate the observed correlation between two variables, the product of a_x, *r*A, and a_y would be taken. Similar calculations can be made to determine the extent to which the shared environment and the nonshared environment contribute to the correlation between the variables. Summing the genetic, shared environmental, and nonshared environmental contributions recreates the observed correlation. Thus, these pathways are somewhat akin to a mediational process as the covariance between the variables is mediated through the latent ACE factors. We will make use of this property to calculate the proportion of the observed correlation that can be attributable to genetic, shared environmental, and nonshared environmental factors.

In addition to behavioral genetic models, we also fit longitudinal *cross-lagged* models, which draw on temporal orderings to make inferences regarding the

directionalities, of effects. This type of model is represented in Figure 11.C. Cross-lagged models are composed of a number of distinct paths. Autoregressive paths reflect the stability of the same variable across time (parameters b_1 and b_4). Cross-paths lead from a predictor variable at one point in time to a different outcome variable at a later point in time (parameters b_2 and b_3). Significant cross-paths indicate a time-ordered relation between two variables while controlling for stability in each variable. This can establish the directionality of effects. Additionally, within-wave (residual) correlations between each variable are estimated. Cross-lagged models were conducted with one variable from each domain (academic behaviors, academic achievement, and parental expectations) resulting in eight separate path models for ECLS-B to incorporate both parent and teacher report and four separate path models for ECLS-K. Additionally, we fit each model with and without control variables to determine the sensitivity of the results.

The transactional model predicts that children will influence their own development by way of influencing their parent and vice versa. To test this hypothesis, we will calculate the total indirect effect that originates from the child through the parent and ending in the child. We will calculate the similar pathway originating from the parent. Mediation models (Baron & Kenny, 1986) were used to evaluate whether some portion of the influence of educational expectations on child development can be accounted for by the proximal parenting behaviors of involvement and stimulation. To test whether the identified processes differ as a function of sociodemographics, we fit multi-group structural equation models. We compared a model in which the focal parameters are allowed to be free across sociodemographic groups (e.g., socioeconomic

status, minority status, and gender) and a model that constrains the parameters to be equal across groups. Because ECLS-K includes more waves of data, it is more likely that interpretable and consistent patterns will emerge in this dataset. Therefore, we focus the moderation analysis on ECLS-K. We use differences in χ^2 and the comparative fit index (CFI) to compare the fit of the models to the data. If the model with all parameters constrained fits significantly worse than the model with all parameters free, then this indicates that the transactional process differs on the basis of sociodemograpics.

All analyses were conducted with M*plus* statistical software using full-information maximum-likelihood estimation to account for missing data (Muthén & Muthén, 1998-2010). To avoid gender differences distorting parameter estimates in our behavior genetic models, we residualized of the influence of gender and standardized all variables (McGue & Bouchard, 1984). If this is uncontrolled for, it has the effect of inflating estimates of heritability because dizygotic, but not monozygotic, twins can have opposite sexes. For analyses using the full ECLS samples, the complex survey option of M*plus* was implemented to weight the results to be representative of the population, and the cluster option was used to account for nonindependence of students sampled from the same sampling frame.

RESULTS

How are educational expectations, academic behaviors and academic outcomes related?

Table 18 presents the correlation matrix for all variables taken from ECLS-B. Significant correlations are found in the expected direction for all variables. That is, educational expectations correlate positively with achievement (*r*'s range from .11 to .17) and parent report of approaches towards learning (*r*'s range from .16 to .21) and negatively with problem behavior (*r*'s range from -.12 to -.15). Results are similar for teacher report of approaches towards learning (*r*'s range from .10 to .13) and problem behavior (*r*'s range from -.05 to -.10). Parent report of approaches towards learning correlate positively with achievement (*r*'s range from .23 to .31), and problem behavior correlates negatively with achievement (*r*'s range from -.15 to -.20). Results are similar for teacher report of approaches towards learning (*r*'s range from .25 to .34) and problem behavior (*r*'s range from -.15 to -.19). Within domains, approaches towards learning are only moderately related to problem behaviors (*r*'s ranging from -.31 to -.40) for parent report, and the results are similar for teacher report (*r*'s range from -.33 to -.59). The association between math and reading achievement was much stronger (*r*'s ranging from .65 to .81). Despite the fact that many of the associations between expectations, academic behaviors and achievement are small to moderate, they are impressive in the sense that they exist even prior to kindergarten entry.

Table 18. Phenotypic Correlations for all ECLS-B Variables

Measure	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
1. Educational Expectations 4	1													
2. Educational Expectations K	.59	1												
3. Math 4	.17	.13	1											
4. Math K	.15	.15	.72	1										
5. Reading 4	.15	.11	.76	.65	1									
6. Reading K	.16	.15	.66	.81	.66	1								
7. Parent Approach Towards Learning 4	.21	.16	.28	.25	.28	.23	1							
8. Parent Approach Towards Learning K	.16	.20	.28	.31	.26	.27	.55	1						
9. Parent Problem Behavior 4	-.14	-.12	-.20	-.18	-.20	-.16	-.40	-.31	1					
10. Parent Problem Behavior K	-.12	-.15	-.16	-.16	-.16	-.15	-.31	-.37	.64	1				
11. Teacher Approach Towards Learning 4	.13	.10	.29	.28	.27	.25	.26	.25	-.23	-.22	1			
12. Teacher Approach Towards Learning K	.11	.11	.32	.34	.30	.30	.24	.34	-.25	-.29	.42	1		
13. Teacher Problem Behavior 4	-.08	-.05	-.19	-.15	-.18	-.15	-.19	-.20	.30	.32	-.50	-.33	1	
14. Teacher Problem Behavior K	-.09	-.10	-.18	-.18	-.17	-.19	-.20	-.29	.33	.37	-.35	-.59	.46	1

Note. All $p < .05$. Variables labeled with 4 refer to the age 4 wave, and variables marked with K refer to the kindergarten wave.

Stability coefficients for each variable are generally high (*r*'s range from .42 to .72). Convergent validity between parent and teacher report of attitudes towards learning and problem behaviors was modest (*r*'s range from .24 to .37). We focus the remainder of our analyses on parent reports, as larger sample sizes were available for parent reports than for teacher reports.

Are parental educational expectations associated with children's genes?

Figure 12 presents the results of a univariate behavior genetic decomposition of each outcome variable at the age 4 and kindergarten waves of ECLS-B. The results are presented graphically in terms of proportion of variance accounted for by genetic or environmental effects. Each estimated proportion of variance is significantly different from zero at $p < .001$ with the exception of the shared environmental parameter for parent and teacher report of approaches towards learning. Nonsignificant variance components were dropped from later models to facilitate convergence. All models fit the data well with no significant χ^2 estimates of misfit.

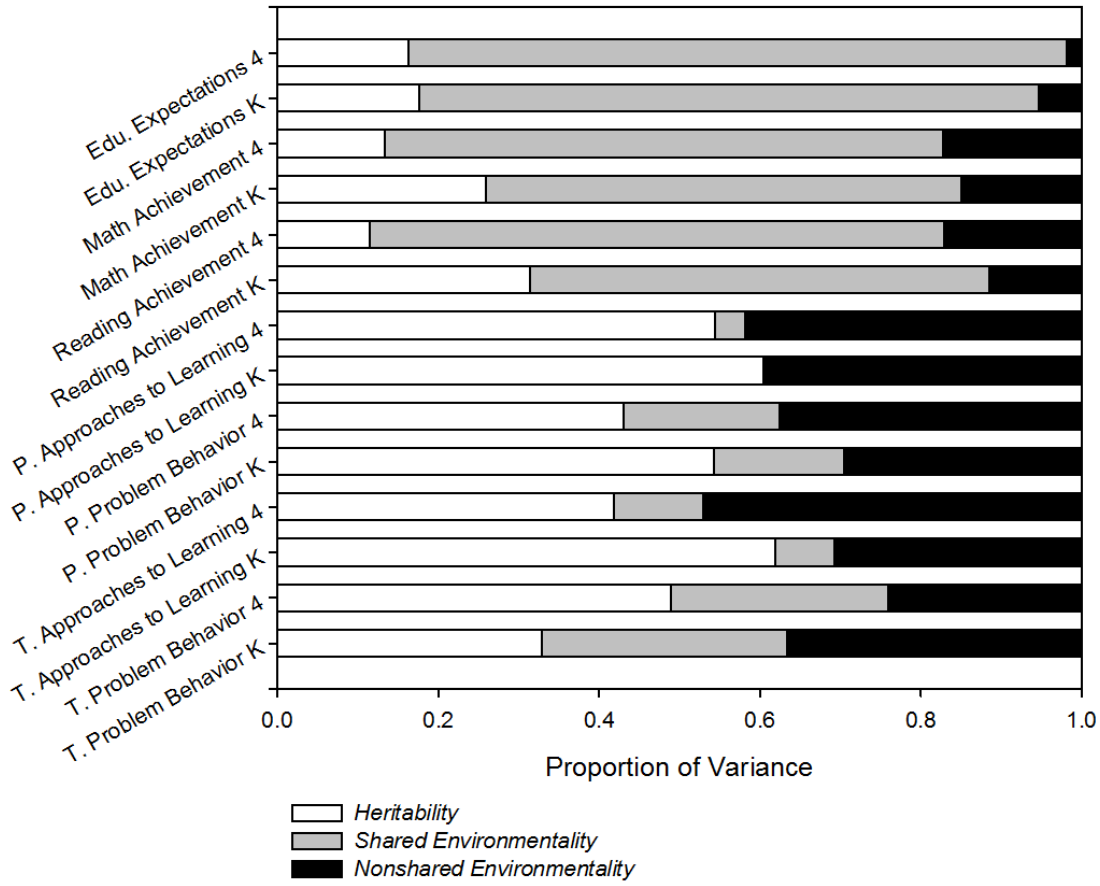

Figure 12. Graphical representation of results from the univariate ACE decomposition for each outcome taken from ECLS-B. Bars represent proportion of variance in the outcome attributable to additive genetic effects, shared environmental effects and nonshared environmental effects.

Parental educational expectations display significant child genetic influence at both time points. Approximately 17% of the variance in educational expectations was related to genotypic differences between children. This indicates that parents are responsive to genetically influenced differences in their children, or that children even as

young as 4 years old are engaged in actively shaping their parent's expectations. However, the variance in educational expectations can primarily be attributed to the shared environment. Roughly three quarters of variation in parental educational expectations can be attributed to between-family variation, pointing to the importance of parents for generating academic beliefs or the influence of structural constraints (e.g., socioeconomic status). The nonshared environment, representing within-family variation and measurement error, accounted for very little variation in expectations, but the estimate was still significant. These results indicate that parental educational expectations are partly influenced by characteristics of the child.

Turning to the child characteristics, each achievement outcome displayed a similar pattern of small, but significant, genetic influence and large shared environmental influence. Approximately 20% of the variance in achievement could be attributed to genotypic differences, 65% to shared environmental differences, and the remaining 15% to unique environmental experiences and measurement error. This distribution of variance components is highly congruent with previous work examining the developmental behavior genetics of cognition (Briley & Tucker-Drob, 2013). The academic behavior variables, on the other hand, display relatively large genetic influences and small or nonexistent shared environmental effects. This distribution of variance components is highly congruent with previous work examining personality development (Bouchard & Loehlin, 2001).

What explains the large shared environmental contributions?

To evaluate whether sociodemographics could account for the large estimates of the shared environment for expectations and achievement, we incorporated these variables into our behavior genetic model. Because these variables are somewhat correlated, we entered socioeconomic status into the model to determine its specific effect, and then sequentially added race/ethnicity, pre-school type, and maternal age. Figure 13 presents the proportion of total variance in expectations and achievement attributable to latent shared environmental effects and measured family-level variables. Race and socioeconomic status accounted for 22% of the variance in expectations and achievement on average and thereby reduced the influence of the shared environment by the same amount. Pre-school type and maternal age accounted for very little remaining variance, approximately 3% on average. About twice as much variance in achievement was explained by measured aspects of the shared environment compared to expectations. In each case, the reduced estimate of the shared environmental influence remained substantial and significant at $p < .001$. This indicates that some family-level influence is operating beyond the well-documented influence of sociodemographics.

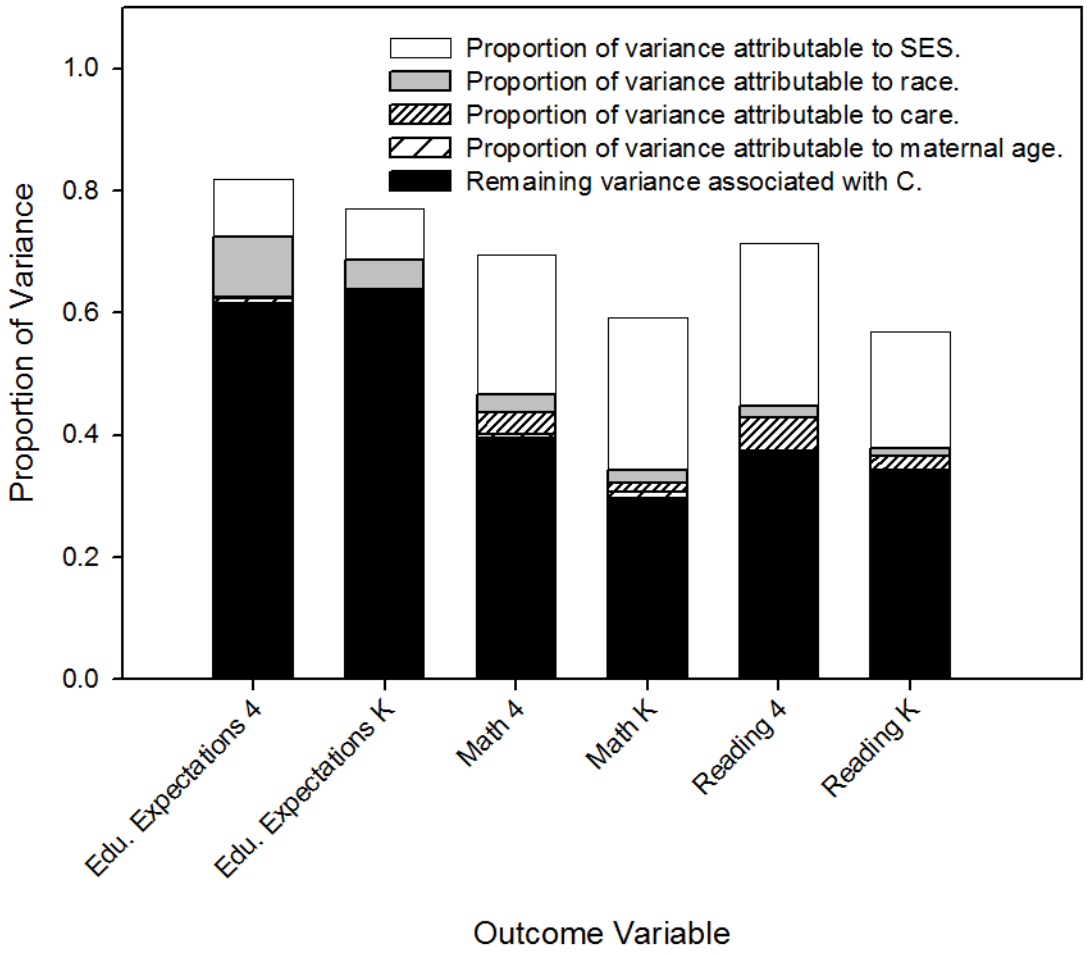

Figure 13. Graphical representation of results from attempts to explain latent C influences with measured family-level environmental variables.

What accounts for the convergent validity of parent and teacher reports?

Parent and teacher reports of approaches towards learning and problem behavior were similarly influenced by genetic, shared environmental, and nonshared environmental factors. However, it may be the case that parents and teachers are rating

different aspects of the children due to the different social contexts in which interaction occurs. To evaluate this possibility, we fit the correlated factors model presented in Figure 11.B to each combination of time points and measures. All models fit the data well (all χ^2 estimates $p > .05$). The results of this analysis are presented in Table 19. Significant genetic correlations were found for both outcomes for every combination of time points. This indicates that parents and teachers are responding, in part, to similar genetic predispositions of the focal child in their rating. There were a few significant nonshared environmental correlations which indicate that parents and teachers agree on which member of a twin pair is higher or lower on the variable controlling for genetic and shared environmental influences. Both of these findings support the validity of parent reports of child behavior and reduce concerns about report bias obscuring results. For problem behaviors, negative correlations were observed between shared environmental factors, but the standard errors were very large for each estimate. Were family background characteristics controlled, this indicates that the convergent validity of parent and teacher report would be higher. Examining the proportion of the observed convergent validity due to ACE makes this point clear. In each case, the majority of the phenotypic correlation is due to genetic factors. For problem behavior, the proportion due to A actually exceeds 1 in some instances due to the countervailing C influence.

Table 19. Convergent ACE Validity of Parent Report with Teacher Report of Approaches Towards Learning and Problem Behavior

Panel 1	Attitudes Towards Learning		Problem Behavior		
Time Combination	*r*A	*r*E	*r*A	*r*C	*r*E
1. Pre-K with Pre-K	.33 (.08) ***	.17 (.08) *	.65 (.17) ***	-.28 (.38)	.24 (.10) *
2. K with K	.31 (.07) ***	.30 (.09) **	.75 (.23) **	-.33 (.45)	.14 (.11)
3. Pre-K with K	.32 (.07) ***	.05 (.09)	.99 (.32) **	-.45 (.39)	-.05 (.10)
4. K with Pre-K	.25 (.09) **	.26 (.10) *	.35 (.17) *	-.02 (.37)	.20 (.12)
Panel 2	Attitudes Towards Learning		Problem Behavior		
Time Combination	prop. due to A	prop. due to E	prop. due to A	prop. due to C	prop. due to E
1. Pre-K with Pre-K	.73	.27	.98	-.22	.24
2. K with K	.67	.33	1.07	-.23	.16
3. Pre-K with K	.92	.08	1.51	-.45	-.06
4. K with Pre-K	.58	.42	.78	-.02	.24

Note. Panel 1 reports the correlation between genetic, shared environmental, and nonshared environmental influences on parent and teacher report. Panel 2 reports the proportion of the observed phenotypic correlation that can be attributed to genetic, shared environmental, and nonshared environmental influences. The timing of parent report is represented by the first time point, and the timing of teacher report is represented with the second time point. The proportion due to a variance component can be greater than 1 if the direction of correlation differs across variance components (e.g., positive *r*A but negative *r*C).
* $p < .05$; ** $p < .01$; *** $p < .001$

What mechanisms account for the stability of the outcomes?

Table 20 presents the genetic and environmental correlations and proportions of the phenotypic correlations due to genetic and environmental factors for the stability of the study outcomes. All models fit the data well (χ^2 estimates $p > .05$). Significant stability of genetic effects was found for each outcome. Stability of shared environmental effects was found for expectations and achievement, but not problem behaviors. Nonshared environmental effects were also moderately stable for each outcome except problem behavior. Nearly the entire stability of expectations was due to the shared environment (95%). For achievement, this was also the case, but genetic effects were three times as influential (15%) compared to genetic influences on stability of expectations. Finally, stability of child academic behavior was primarily due to genetic factors (57% to 83%), but both shared and nonshared environmental factors contributed to a lesser extent.

Table 20. ACE stability and proportions of stability due to ACE for study outcomes

	*r*A	*r*C	*r*E	prop. A	prop. C	prop. E
1. Educational Expectations	.17 (.08) *	.74 (.03) ***	.08 (.09)	.05	.95	.00
2. Math	.56 (.16) ***	.90 (.04) ***	.39 (.07) ***	.15	.77	.08
3. Reading	.50 (.17) **	.85 (.05) ***	.22 (.08) **	.15	.81	.04
4. Parent ATL	.79 (.05) ***	-	.24 (.07) **	.83	-	.17
5. Parent Problem Behavior	.99 (.32) **	-.45 (.39)	-.05 (.10)	.59	.20	.21
6. Teacher ATL	.34 (.09) ***	-	.31 (.11) **	.65	-	.35
7. Teacher Problem Behavior	.65 (.27) *	.37 (.27)	.22 (.14)	.57	.27	.16

Note. ATL stands for approaches towards learning. Prop. stands for proportion of the observed phenotypic correlations due to the genetic or environmental variance component.
* $p < .05$; ** $p < .01$; *** $p < .001$

What mechanisms link parental educational expectations and child factors?

Table 21 presents the genetic and environmental correlations between educational expectations and the child factors within and across waves. Each model fit the data well (χ^2 estimates $p > .05$) or adequately (two models: χ^2 estimates $p > .01$ with CFI values of .98 and .99). It is likely that the χ^2 test is overpowered for these two exceptions as other indicators of fit are excellent. Table 22 presents the proportion of the observed phenotypic correlation that is due to genetic or environmental factors. The full table of genetic and environmental correlations is available in Table 23.

Table 21. Genetic and Environmental Correlates of Educational Expectations

Panel 1: Within Wave	Age 4 Expectations			Kindergarten Expectations		
Child Factors	*r*A	*r*C	*r*E	*r*A	*r*C	*r*E
1. Math	.22 (.12)	.22 (.05) ***	.12 (.08)	.34 (.12) **	.17 (.07) **	.12 (.08)
2. Reading	.33 (.15) *	.20 (.05) ***	-.04 (.08)	.14 (.11)	.21 (.07) **	.15 (.08)
3. Parent ATL	.42 (.08) ***	-	-.09 (.08)	.39 (.09) ***	-	.14 (.08)
4. Parent Problem Behavior	-.14 (.10)	-.22 (.11) *	.07 (.08)	-.26 (.12) *	-.19 (.14)	.11 (.08)
5. Teacher ATL	.19 (.12)	-	.09 (.12)	.21 (.11)	-	-.06 (.12)
6. Teacher Problem Behavior	-.24 (.13)	.04 (.12)	-.10 (.12)	-.38 (.23)	.09 (.13)	.08 (.13)
Panel 2: Across Wave	Age 4 Expectations			Kindergarten Expectations		
Child Factors	*r*A	*r*C	*r*E	*r*A	*r*C	*r*E
1. Math	.27 (.10) **	.24 (.06) ***	-.02 (.09)	.03 (.17)	.31 (.06) ***	.09 (.08)
2. Reading	.19 (.09) *	.23 (.06) ***	-.04 (.09)	.05 (.18)	.21 (.06) ***	.14 (.08)
3. Parent ATL	.50 (.10) ***	-	-.21 (.09) *	.45 (.12) ***	-	-.24 (.08) **
4. Parent Problem Behavior	.07 (.10)	-.32 (.14) *	-.02 (.09)	-.24 (.14)	-.19 (.13)	.14 (.08)
5. Teacher ATL	.27 (.10) **	-	-.08 (.09)	.43 (.15) **	-	.11 (.13)
6. Teacher Problem Behavior	-.30 (.16)	-.02 (.11)	.02 (.09)	-.32 (.17)	-.02 (.14)	.09 (.13)

Note. ATL stands for approaches towards learning. Standardized parameter estimates are presented first, followed by standard errors in parentheses. Panel 1 represents within wave correlations between expectations and child factors. Panel 2 represents across wave correlations between expectations and child factors.
* $p < .05$; ** $p < .01$; *** $p < .001$

Table 22. Proportion of the Observed Correlation between Educational Expectations and Child Factors due to ACE

Panel 1: Within Wave	Age 4 Expectations			Kindergarten Expectations		
Child Factors	prop. A	prop. C	prop. E	prop. A	prop. C	prop. E
1. Math	.16	.81	.03	.36	.58	.06
2. Reading	.23	.78	-.01	.18	.76	.06
3. Parent ATL	1.07	-	-.07	.87	-	.13
4. Parent Problem Behavior	.31	.74	-.05	.60	.51	-.11
5. Teacher ATL	.88	-	.12	1.11	-	-.11
6. Teacher Problem Behavior	1.19	-.31	.12	2.44	-1.16	-.28
Panel 2: Across Wave	Age 4 Expectations			Kindergarten Expectations		
Child Factors	prop. A	prop. C	prop. E	prop. A	prop. C	prop. E
1. Math	.25	.75	.00	.02	.94	.04
2. Reading	.21	.80	-.01	.04	.88	.08
3. Parent ATL	1.14	-	-.14	1.35	-	-.35
4. Parent Problem Behavior	.14	.85	.01	.56	.60	-.16
5. Teacher ATL	1.07	-	-.07	.88	-	.12
6. Teacher Problem Behavior	.92	.11	-.03	.99	.12	-.11

Note: ATL stands for approaches towards learning. Prop. stands for proportion of the observed phenotypic correlations due to the genetic or environmental variance component. Panel 1 represents within wave proportions for expectations and child factors. Panel 2 represents across wave proportions between expectations and child factors. The proportion due to a variance component can be greater than 1 if the direction of correlation differs across variance components (e.g., positive *r*A but negative *r*C).

Table 23. Full Genetic and Environmental Correlations for All ECLS-B Variables

*r*A	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Edu. Expectations Pre-K	1												
2. Edu. Expectations K	.17 (.08) *	1											
3. Math Pre-K	.22 (.12)	.03 (.17)	1										
4. Math K	.27 (.10) **	.34 (.12) **	.56 (.16) ***	1									
5. Reading Pre-K	.33 (.15) *	.05 (.18)	.89 (.25) ***	.56 (.20) **	1								
6. Reading K	.19 (.09) *	.14 (.11)	.58 (.15) ***	.96 (.11) ***	.50 (.17) **	1							
7. Parent ATL Pre-K	.42 (.08) ***	.45 (.12) ***	.94 (.20) ***	.78 (.12) ***	1.20 (.28) ***	.68 (.10) ***	1						
8. Parent ATL K	.50 (.10) ***	.39 (.09) ***	.82 (.18) ***	.64 (.10) ***	.97 (.24) ***	.70 (.10) ***	.79 (.05) ***	1					
9. Parent Prob. Beh. Pre-K	-.14 (.10)	-.24 (.14)	.00 (.23)	-.37 (.18) *	.00 (.25)	-.32 (.15) *	-.40 (.07) ***	-.45 (.09) ***	1				
10. Parent Prob. Beh. K	.07 (.10)	-.26 (.12) *	.02 (.22)	-.40 (.15) **	.09 (.23)	-.50 (.14) ***	-.32 (.08) ***	-.51 (.08) ***	.99 (.32) **	1			

Table 23, cont.

11. Teacher ATL Pre-K	.19 (.12)	.43 (.15) **	.67 (.19) **	.80 (.16) ***	.93 (.26) ***	.73 (.13) ***	.33 (.08) ***	.25 (.09) **	-.32 (.11) **	-.28 (.11) *	1		
12. Teacher ATL K	.27 (.10) **	.21 (.11)	.96 (.22(***	.82 (.12) ***	.89 (.25) ***	.64 (.11) ***	.32 (.07) ***	.31 (.07) ***	-.33 (.10) **	-.30 (.08) ***	.34 (.09) ***	1	
13. Teacher Prob. Beh. Pre-K	-.24 (.13)	-.32 (.17)	-.15 (.25)	-.53 (.20) **	-.45 (.29)	-.37 (.16) *	-.28 (.10) **	-.29 (.11) **	.65 (.17) ***	.35 (.17) *	-.73 (.09) ***	-.25 (.10) *	1
14. Teacher Prob. Beh. K	-.30 (.16)	-.38 (.23)	-.87 (.37) *	-.45 (.24)	-.01 (.36)	-.28 (.21)	-.38 (.15) **	-.44 (.15) **	.99 (.32) **	.75 (.23) **	-.38 (.17) *	-.67 (.11) ***	.65 (.27) *
*r*C	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Edu. Expectations Pre-K	1												
2. Edu. Expectations K	.74 (.03) ***	1											
3. Math Pre-K	.22 (.05) ***	.31 (.06) ***	1										
4. Math K	.24 (.06) ***	.17 (.07) **	.90 (.04) ***	1									
5. Reading Pre-K	.20 (.05) ***	.21 (.06) ***	.89 (.04) ***	.85 (.05) ***	1								
6. Reading K	.23 (.06) ***	.21 (.07) **	.83 (.05) ***	.92 (.04) ***	.85 (.05) ***	1							
7. Parent ATL Pre-K	-	-	-	-	-	-	1						
8. Parent ATL K	-	-	-	-	-	-	-	1					

Table 23, cont.

9. Parent Prob. Beh. Pre-K	-.22 (.11) *	-.19 (.13)	-.44 (.15) **	-.20 (.16)	-.42 (.15) **	-.18 (.16)	-	-	1				
10. Parent Prob. Beh. K	-.32 (.14) *	-.19 (.14)	-.44 (.18) *	-.28 (.17)	-.50 (.19) **	-.15 (.17)	-	-	-.45 (.39)	1			
11. Teacher ATL Pre-K	-	-	-	-	-	-	-	-	-	-	1		
12. Teacher ATL K	-	-	-	-	-	-	-	-	-	-	-	1	
13. Teacher Prob. Beh. Pre-K	.04 (.12)	-.02 (.14)	-.22 (.14)	-.01 (.17)	-.12 (.14)	-.07 (.17)	-	-	-.28 (.38)	-.02 (.37)	-	-	1
14. Teacher Prob. Beh. K	-.02 (.11)	.09 (.13)	-.10 (.14)	-.25 (.15)	-.33 (.15) *	-.32 (.16) *	-	-	-.45 (.39)	-.33 (.45)	-	-	.37 (.27)
*r*E	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Edu. Expectations Pre-K	1												
2. Edu. Expectations K	.08 (.09)	1											
3. Math Pre-K	.12 (.08)	.09 (.08)	1										
4. Math K	-.02 (.09)	.12 (.08)	.39 (.07) ***	1									
5. Reading Pre-K	-.04 (.08)	.14 (.08)	.15 (.08)	.17 (.08) *	1								
6. Reading K	-.04 (.09)	.15 (.08)	.19 (.08) *	.18 (.08) *	.22 (.08) **	1							

Table 23, cont.

7. Parent ATL Pre-K	-.09 (.08)	-.24 (.08) **	.01 (.07)	.06 (.08)	-.09 (.07)	-.10 (.08)	1						
8. Parent ATL K	-.21 (.09) *	.14 (.08)	.15 (.08) *	.26 (.08) **	.01 (.08)	.01 (.08)	.24 (.07) **	1					
9. Parent Prob. Beh. Pre-K	.07 (.08)	.14 (.08)	-.14 (.08)	.01 (.08)	-.08 (.08)	-.04 (.08)	-.41 (.06) ***	-.21 (.07) **	1				
10. Parent Prob. Beh. K	-.02 (.09)	.11 (.08)	-.14 (.09)	.04 (.09)	-.09 (.09)	.14 (.08)	-.24 (.08) **	-.24 (.07) **	-.05 (.10)	1			
11. Teacher ATL Pre-K	.09 (.12)	.11 (.13)	.13 (.09)	.05 (.12)	-.02 (.09)	-.12 (.10)	.17 (.08) *	.26 (.10) *	-.19 (.08) *	-.10 (.12)	1		
12. Teacher ATL K	-.08 (.09)	-.06 (.12)	-.10 (.09)	.01 (.10)	-.05 (.10)	-.15 (.10)	.05 (.09)	.30 (.09) **	-.10 (.09)	-.08 (.10)	.31 (.11) **	1	
13. Teacher Prob. Beh. Pre-K	-.10 (.12)	.09 (.13)	-.14 (.11)	.09 (.12)	.10 (.10)	.09 (.10)	-.15 (.09)	-.13 (.10)	.24 (.10) *	.20 (.12)	-.40 (.08) ***	-.14 (.13)	1
14. Teacher Prob. Beh. K	.02 (.09)	.08 (.13)	.11 (.10)	.00 (.11)	.01 (.11)	.16 (.11)	-.02 (.09)	-.20 (.09) *	-.05 (.10)	.14 (.11)	-.22 (.12)	-.47 (.08) ***	.22 (.14)

Note. ATL stands for approaches towards learning.
* $p < .05$; ** $p < .01$; *** $p < .001$

As shown in Table 21, there were significant shared environmental correlations between parental expectations and child achievement in reading and math. This result indicates that parents who had higher expectations for both their children, on average, had higher achieving children. Put differently, between-family differences in educational expectations were correlated with between-family differences in achievement. This was true at both age 4 and at kindergarten, both within-waves and across-waves. In addition, there were shared environmental correlations between educational expectations and problem behavior, but only for age 4 expectations: Parents who perceived their children to have less problem behavior had higher expectations, on average, for their children.

There were also genetic correlations between child characteristics and parental expectations. With regards to achievement (math and reading), four of the eight possible genetic correlations between achievement and educational expectations were statistically significant. The largest genetic correlations were between reading at age 4 and expectations at age 4 (.33) and between math at kindergarten and expectations at kindergarten (.34). In addition, there was a consistent genetic correlation between approaches towards learning and expectations; this was evident at age-4 and at kindergarten, both within- and across-waves. These genetic correlations indicate that within-family variation in educational expectations is associated with genetic differences in measured child characteristics, particularly approaches towards learning.

Do transactional processes occur at this very early stage of development?

We performed a series of longitudinal cross-lagged path models using data from the entire ECLS-B sample to clarify the longitudinal processes that link child characteristics and parental expectations. Because we used four indicators each of child academic behaviors and two indicators of child academic achievement, we fit eight trivariate models based on combinations of key study variables. Additionally, we report results with and without control variables included in the model. As the two-wave cross-lagged path model is fully saturated, it has perfect fit to the data. We will primarily focus on the model including parent-report of approaches towards learning and math achievement. Figure 14 presents the standardized parameter estimates from this model. Estimates listed first come from a model that did not include control variables. The second estimates come from a model that additionally controlled for variation in family socioeconomic status, race/ethnicity, child gender, maternal age, pre-school care arrangement, and disability status. The results were largely unchanged by the addition of the controls.

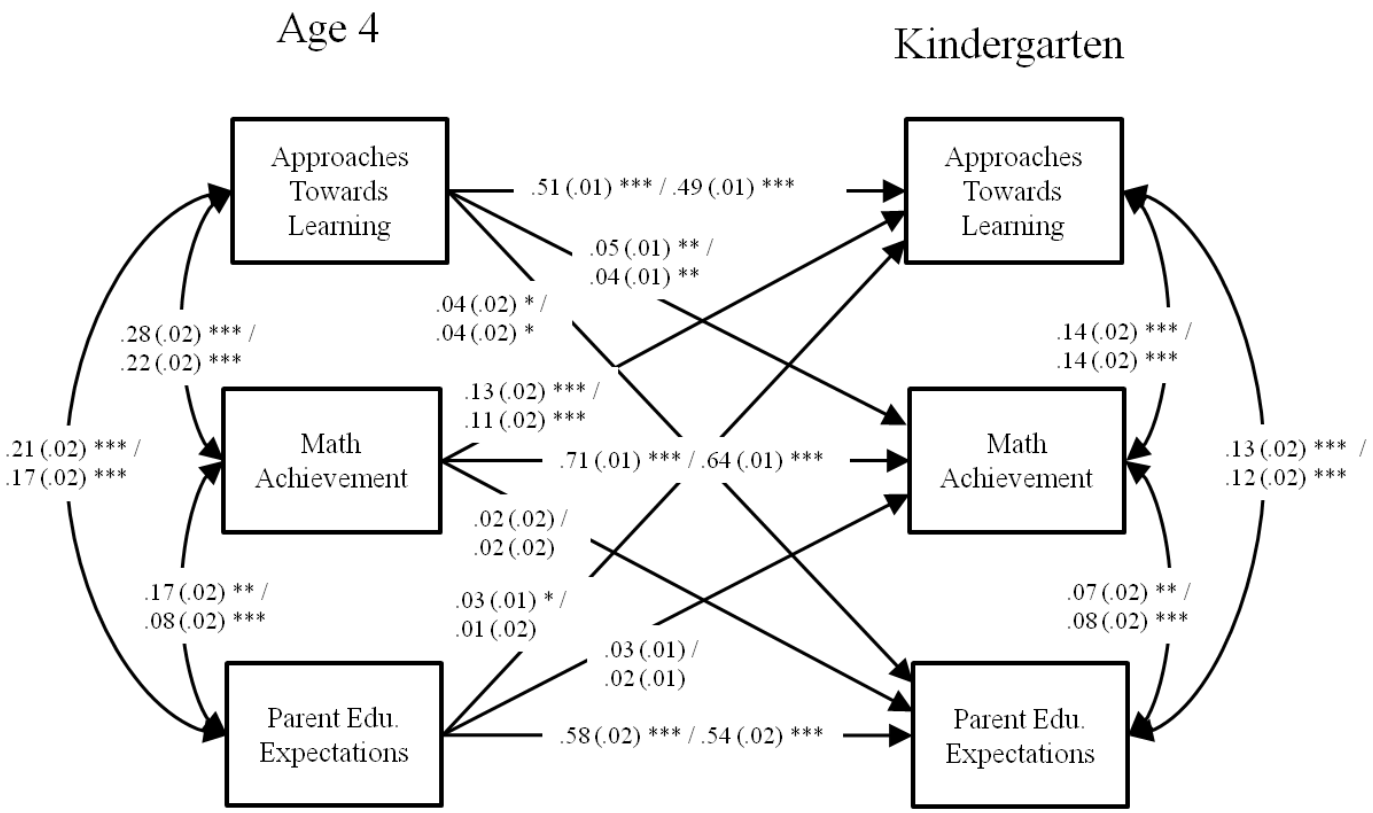

Figure 14. Cross-lagged path model using ECLS-B data and the variables of approaches towards learning, math ability, and educational expectations. Standardized parameters are reported first, followed by standard errors and significance levels. Uncontrolled results are presented first followed by results from a model that included controls. * $p < .05$; ** $p < .01$; *** $p < .001$.

Several pathways are of note. First, there was substantial stability for each outcome. High stability places an upper limit on the amount of transactional effects that can be observed over the interval of a year. Second, higher levels of approaches towards learning at the initial time point predict higher levels of achievement and expectations at the second time point. Third, earlier math achievement predicts higher approaches towards learning at the later time point, but not expectations. Fourth, early expectations have marginally significant relations with later achievement in both uncontrolled and controlled models (p's $< .10$). Fifth, early expectations significantly predicted later approaches towards learning only in the uncontrolled model. Finally, the three variables remained correlated at the second time point.

The full parameter estimates, standard errors, (residual) correlations, and significance levels for the remaining models can be found in Tables 24-27. Briefly, these models provide similar levels of support for the transactional model. In general, early approaches towards learning and problem behavior (when parent-reported) predicted later expectations, but early achievement did not. Reciprocally, early expectations positively predicted later approaches towards learning and achievement. However, the association with approaches towards learning tends to be attenuated by controls or the use of teacher report. Overall, there was consilience between the results from the longitudinal model and the behavioral genetic model. For both methodologies, there was the strongest evidence for transactions between child academic behavior – particularly approaches to learning – and parental expectations: Approaches towards learning showed the strongest

genetic correlations both between and across waves, and the strongest cross-lagged paths with subsequent expectations

Table 24. Full standardized parameter estimates of four separate cross-lagged path models from ECLS-B without control variables (Parent report of academic behavior)

Model 1 – Approaches Towards Learning, Math Achievement and Expectations			
		Kindergarten Variable	
Age 4 Variable	Math	ATL	Expectations
Auto-Regressive	.71 (.01) ***	.51 (.01) ***	.58 (.02) ***
Cross-Paths			
Math →		.13 (.02) ***	.02 (.02)
ATL →	.05 (.01) **		.04 (.02) *
Expectations →	.03 (.01)	.03 (.01) *	
(Residual) Correlations			
Math		.28 (.02) ***	.17 (.02) ***
ATL	.14 (.02) ***		.21 (.02) ***
Expectations	.07 (.02) **	.13 (.02) ***	
Model 2 – Problem Behavior, Math Achievement and Expectations			
		Kindergarten Variable	
Age 4 Variable	Math	Problem	Expectations
Auto-Regressive	.71 (.01) ***	.63 (.01) ***	.58 (.02) ***
Cross-Paths			
Math →		-.03 (.02)	.03 (.02)
Problem →	-.03 (.01)		-.04 (.02) *
Expectations →	.03 (.01) *	-.03 (.02)	
(Residual) Correlations			
Math		-.20 (.02) ***	.17 (.02) ***
Problem	-.04 (.02) *		-.14 (.02) ***
Expectations	.07 (.02) ***	-.09 (.02) ***	
Model 3 – Approaches Towards Learning, Reading Achievement and Expectations			
		Kindergarten Variable	
Age 4 Variable	Reading	ATL	Expectations
Auto-Regressive	.64 (.01) ***	.51 (.01) ***	.58 (.02) ***
Cross-Paths			
Reading →		.11 (.02) ***	.01 (.02)
ATL →	.04 (.02) *		.04 (.02) *

Table 24, cont.

Expectations →	.06 (.02) **	.03 (.01) *	
(Residual Correlations)			
Reading		.28 (.02) ***	.15 (.02) ***
ATL	.12 (.02) ***		.21 (.02) ***
Expectations	.07 (.02) **	.13 (.02) ***	

Model 4 – Problem Behavior, Reading Achievement and Expectations

	Kindergarten Variable		
Age 4 Variable	Reading	Problem	Expectations
Auto-Regressive	.65 (.01) ***	.63 (.01) ***	.58 (.02) ***
Cross-Paths			
Reading →		-.03 (.02)	.01 (.02)
Problem →	-.03 (.02)		-.04 (.02) **
Expectations →	.06 (.02) **	-.03 (.02)	
(Residual) Correlations			
Reading		-.20 (.02) ***	.15 (.02) ***
Problem	-.05 (.02) *		-.14 (.02) ***
Expectations	.07 (.02) **	-.09 (.02) ***	

Note. All parameter estimates are standardized with standard errors in parentheses. Correlations listed above the diagonal represent associations at the initial time point, and correlations below the diagonal represent residual correlations at the second time point. ATL refers to approaches towards learning.
* $p < .05$; ** $p < .01$; *** $p < .001$

Table 25. Full standardized parameter estimates of four separate cross-lagged path models from ECLS-B with control variables (Parent report of academic behaviors)

Model 1 – Parent Approaches Towards Learning, Math Achievement and Expectations			
		Kindergarten Variable	
Age 4 Variable	Math	ATL	Expectations
Auto-Regressive	.64 (.01) ***	.49 (.01) ***	.54 (.02) ***
Cross-Paths			
Math →		.11 (.02) ***	.02 (.02)
ATL →	.04 (.01) **		.04 (.02) *
Expectations →	.02 (.01)	.01 (.02)	
(Residual) Correlations			
Math		.22 (.02) ***	.08 (.02) ***
ATL	.14 (.02) ***		.17 (.02) ***
Expectations	.08 (.02) ***	.12 (.02) ***	
Model 2 – Parent Problem Behavior, Math Achievement and Expectations			
		Kindergarten Variable	
Age 4 Variable	Math	Problem	Expectations
Auto-Regressive	.65 (.01) ***	.61 (.01) ***	.54 (.02) ***
Cross-Paths			
Math →		-.03 (.02)	.02 (.02)
Problem →	-.02 (.01)		-.04 (.02) **
Expectations →	.03 (.01) *	-.02 (.02)	
(Residual) Correlations			
Math		-.14 (.02) ***	.08 (.02) ***
Problem	-.04 (.02) **		-.10 (.02) ***
Expectations	.08 (.02) ***	-.08 (.02) ***	
Model 3 – Parent Approaches Towards Learning, Reading Achievement and Expectations			
		Kindergarten Variable	
Age 4 Variable	Reading	ATL	Expectations
Auto-Regressive	.60 (.02) ***	.50 (.01) ***	.54 (.02)
Cross-Paths			
Reading →		.08 (.02) ***	.01 (.01)
ATL →	.03 (.01) *		.04 (.02) *

Table 25, cont.

Expectations →	.04 (.02) *	.01 (.02)	
(Residual Correlations)			
Reading		.22 (.02) ***	.08 (.02) ***
ATL	.11 (.02) ***		.17 (.02) ***
Expectations	.07 (.02) **	.12 (.02) ***	

Model 4 – Parent Problem Behavior, Reading Achievement and Expectations

	Kindergarten Variable		
Age 4 Variable	Reading	Problem	Expectations
Auto-Regressive	.60 (.02) ***	.61 (.01) ***	.54 (.02) ***
Cross-Paths			
Reading →		-.03 (.02)	.01 (.02)
Problem →	-.02 (.02)		-.04 (.02) *
Expectations →	.04 (.02) *	-.03 (.02)	
(Residual) Correlations			
Reading		-.13 (.02) ***	.08 (.02) ***
Problem	-.05 (.02) *		-.10 (.02) ***
Expectations	.07 (.02) **	-.08 (.02) ***	

Note. All parameter estimates are standardized with standard errors in parentheses. Correlations listed above the diagonal represent associations at the initial time point, and correlations below the diagonal represent residual correlations at the second time point. ATL refers to approaches towards learning. All models control for maternal age, type of pre-school care, disability status, child gender, and child race.
* $p < .05$; ** $p < .01$; *** $p < .001$

Table 26. Full standardized parameter estimates of four separate cross-lagged path models from ECLS-B without control variables (Teacher report of academic behavior)

Model 1 – Approaches Towards Learning, Math Achievement and Expectations			
	Kindergarten Variable		
Age 4 Variable	Math	ATL	Expectations
Auto-Regressive	.70 (.01) ***	.35 (.02) ***	.58 (.02) ***
Cross-Paths			
Math →		.21 (.02) ***	.03 (.02)
ATL →	.07 (.02) ***		.02 (.02)
Expectations →	.03 (.01) *	.04 (.03)	
(Residual) Correlations			
Math		.30 (.02) ***	.17 (.02) ***
ATL	.15 (.02) ***		.13 (.02) ***
Expectations	.07 (.02) ***	.05 (02) *	
Model 2 – Problem Behavior, Math Achievement and Expectations			
	Kindergarten Variable		
Age 4 Variable	Math	Problem	Expectations
Auto-Regressive	.72 (.01) ***	.44 (.02) ***	.59 (.02) ***
Cross-Paths			
Math →		-.10 (.02) ***	.04 (.02) *
Problem →	-.01 (.02)		.00 (.02)
Expectations →	.03 (.01) *	-.04 (.02)	
(Residual) Correlations			
Math		-.19 (.02) ***	.17 (.02) ***
Problem	-.08 (.02) **		-.08 (.03) **
Expectations	.08 (.02) ***	-.07 (.03) **	
Model 3 – Approaches Towards Learning, Reading Achievement and Expectations			
	Kindergarten Variable		
Age 4 Variable	Reading	ATL	Expectations
Auto-Regressive	.63 (.01) ***	.37 (.02) ***	.59 (.02) ***
Cross-Paths			
Reading →		.19 (.02) ***	.01 (.02)
ATL →	.07 (.02) ***		.02 (.02)

Table 26, cont.

Expectations →	.06 (.02) **	.04 (.03)	
(Residual Correlations)			
Reading		.27 (.02) ***	.16 (.02) ***
ATL	.12 (.02) ***		.13 (.02) ***
Expectations	.07 (.02) **	.06 (.02) *	

Model 4 – Problem Behavior, Reading Achievement and Expectations

	Kindergarten Variable		
Age 4 Variable	Reading	Problem	Expectations
Auto-Regressive	.65 (.01) ***	.45 (.02) ***	.59 (.02) ***
Cross-Paths			
Reading →		-.09 (.02) ***	.02 (.02)
Problem →	-.03 (.02)		.00 (.02)
Expectations →	.06 (.02) **	-.04 (.02)	
(Residual) Correlations			
Reading		-.18 (.02) ***	.15 (.02) ***
Problem	-.10 (.03) ***		-.08 (.03) **
Expectations	.08 (.02) **	-.07 (.03) **	

Note. All parameter estimates are standardized with standard errors in parentheses. Correlations listed above the diagonal represent associations at the initial time point, and correlations below the diagonal represent residual correlations at the second time point. ATL refers to approaches towards learning.
* $p < .05$; ** $p < .01$; *** $p < .001$

Table 27. Full standardized parameter estimates of four separate cross-lagged path models from ECLS-B with control variables (Teacher report of academic behaviors)

Model 1 – Teacher Approaches Towards Learning, Math Achievement and Expectations			
		Kindergarten Variable	
Age 4 Variable	Math	ATL	Expectations
Auto-Regressive	.63 (.02) ***	.32 (.03) ***	.54 (.02) ***
Cross-Paths			
Math →		.19 (.03) ***	.02 (.02)
ATL →	.07 (.02) **		.02 (.02)
Expectations →	.02 (.01)	.02 (.02)	
(Residual) Correlations			
Math		.25 (.02) ***	.08 (.02) ***
ATL	.14 (.02) ***		.10 (.02) ***
Expectations	.08 (.02) ***	.04 (.02) ***	
Model 2 – Teacher Problem Behavior, Math Achievement and Expectations			
		Kindergarten Variable	
Age 4 Variable	Math	Problem	Expectations
Auto-Regressive	.65 (.01) ***	.41 (.02) ***	.54 (.02) ***
Cross-Paths			
Math →		-.09 (.02) ***	.03 (.02)
Problem →	.00 (.02)		.00 (.02)
Expectations →	.03 (.01) *	.02 (.02)	
(Residual) Correlations			
Math		-.13 (.03) ***	.08 (.02) ***
Problem	-.07 (.02) **		-.04 (.03)
Expectations	.08 (.02) ***	-.05 (.03) *	
Model 3 – Teacher Approaches Towards Learning, Reading Achievement and Expectations			
		Kindergarten Variable	
Age 4 Variable	Reading	ATL	Expectations
Auto-Regressive	.59 (.02) ***	.33 (.02) ***	.54 (.02) ***
Cross-Paths			
Reading →		.17 (.03) ***	.01 (.02)
ATL →	.06 (.02) ***		.02 (.02)

Table 27, cont.

Expectations →	.04 (.02) *	.03 (.02)	
(Residual Correlations)			
Reading		.21 (.02) ***	.08 (.02) ***
ATL	.12 (.02) ***		.09 (.02) ***
Expectations	.06 (.02) **	.05 (.02) *	

Model 4 – Teacher Problem Behavior, Reading Achievement and Expectations

	Kindergarten Variable		
Age 4 Variable	Reading	Problem	Expectations
Auto-Regressive	.60 (.02) ***	.41 (.02) ***	.54 (.02) ***
Cross-Paths			
Reading →		-.08 (.02) ***	.02 (.02)
Problem →	-.01 (.02)		.00 (.02)
Expectations →	.04 (.02) *	-.02 (.02)	
(Residual) Correlations			
Reading		-.11 (.02) ***	.08 (.02) ***
Problem	-.10 (.03) ***		-.04 (.03)
Expectations	.07 (.02) **	-.05 (.03) *	

Note. All parameter estimates are standardized with standard errors in parentheses. Correlations listed above the diagonal represent associations at the initial time point, and correlations below the diagonal represent residual correlations at the second time point. ATL refers to approaches towards learning. All models control for maternal age, type of pre-school care, disability status, child gender, and child race.
* $p < .05$; ** $p < .01$; *** $p < .001$

How do transactional processes develop as children progress academically?

The ECLS-K data can act as an extension of the previous results and allow an examination of how these reciprocal effects develop as children grow and gain more independence over their environment. Importantly, this analysis is based on a separate dataset containing different individuals, and therefore represents a conceptual extension rather than direct longitudinal follow-up of the same children. Figure 15 presents a similar, trivariate cross-lagged path model beginning in kindergarten and ending in fifth

grade for teacher report of approaches towards learning, math achievement, and educational expectations. The full parameter estimates, standard errors, (residual) correlations and significance levels can be found in Tables 28-31. Each model recaptured the data well as indicated by excellent model fit statistics (CFI = .95-.97). Again, we report results from an uncontrolled model and a model that controls for variation in family socioeconomic status, race/ethnicity, child gender, maternal age, pre-school care arrangement, and disability status.

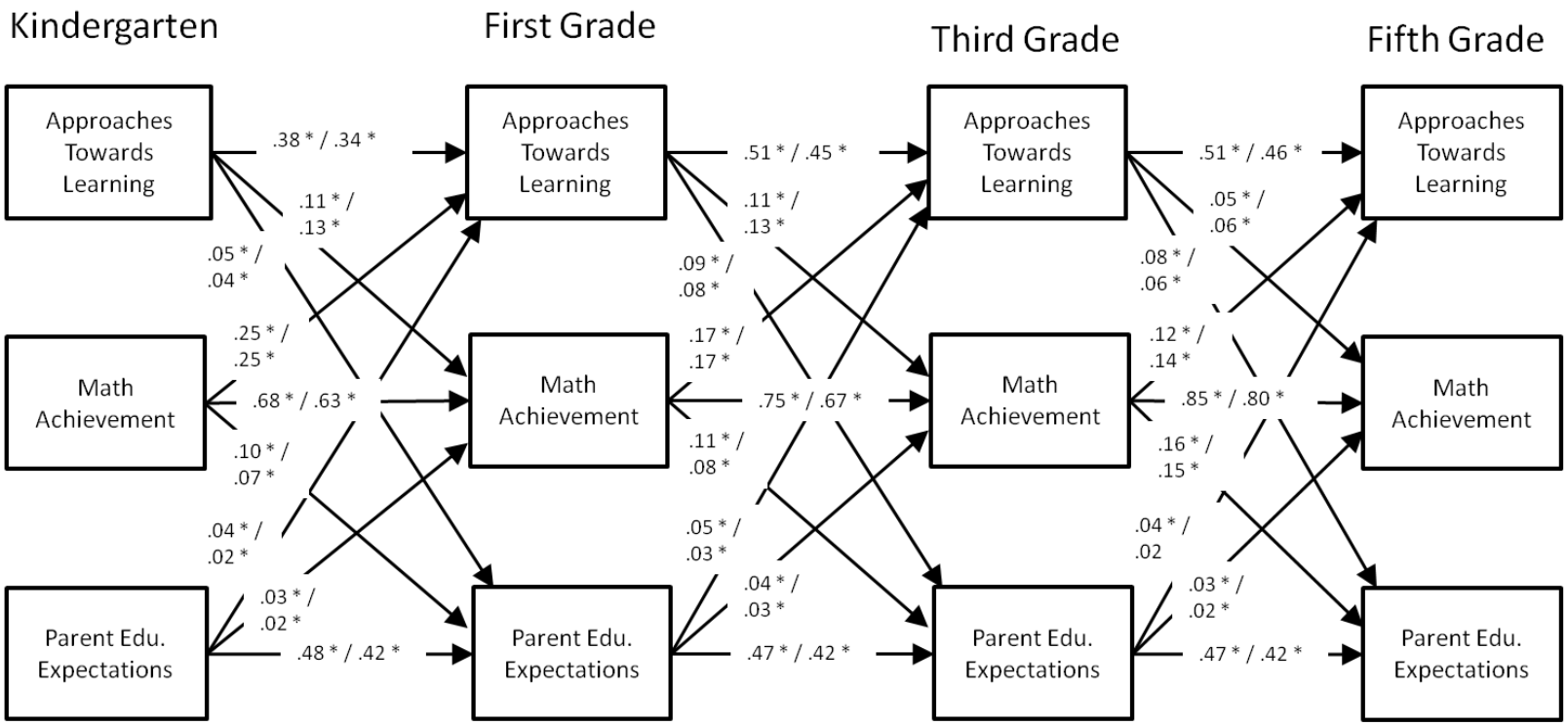

Figure 15. Cross-lagged path model using ECLS-K data and the variables of approaches towards learning, math ability, and educational expectations. Standardized parameters with significance levels are reported (* $p < .05$). Uncontrolled results are presented first followed by results from a model that included controls. To reduce clutter, standard errors and complete significance levels are not reported.

Table 28. Full standardized parameter estimates of four separate cross-lagged path models from ECLS-K without controls

Model 1 – Approaches Towards Learning, Math Achievement and Expectations			
		Next Wave	
Previous Wave	Math	ATL	Expectations
Autoregressive			
Kindergarten	.68 (.01) ***	.38 (.01) ***	.48 (.01) ***
First Grade	.75 (.01) ***	.51 (.01) ***	.47 (.01) ***
Third Grade	.85 (.01) ***	.51 (.01) ***	.47 (.02) ***
Cross-lagged Paths			
Kindergarten			
Math →		.25 (.01) ***	.10 (.01) ***
ATL →	.11 (.01) ***		.05 (.01) ***
Expectations →	.03 (.01) ***	.04 (.01) ***	
First Grade			
Math →		.17 (.01) ***	.11 (.01) ***
ATL →	.11 (.01) ***		.09 (.01) ***
Expectations →	.04 (.01) ***	.05 (.01) ***	
Third Grade			
Math →		.12 (.01) ***	.16 (.01) ***
ATL →	.05 (.01) ***		.08 (.01) ***
Expectations →	.03 (.01) ***	.04 (.01) **	
Model 2 – Problem Behavior, Math Achievement and Expectations			
		Next Wave	
Previous Wave	Math	Problem	Expectations
Autoregressive			
Kindergarten	.72 (.01) ***	.50 (.01) ***	.48 (.01) ***
First Grade	.79 (.01) ***	.54 (.01) ***	.47 (.01) ***
Wave 5	.86 (.01) ***	.53 (.01) ***	.48 (.02) ***
Cross-lagged Paths			
Kindergarten			
Math →		-.08 (.02) ***	.12 (.01) ***
Problem →	-.06 (.01) ***		-.03 (.01) *
Expectations →	.04 (.01) ***	-.02 (.01) *	
First Grade			
Math →		-.07 (.01) ***	.14 (.01) ***

Table 28, cont.

Problem →	-.03 (.01) ***		-.05 (.01) ***
Expectations →	.05 (.01) ***	-.05 (.01) ***	
Third Grade			
Math →		-.04 (.01) ***	.18 (.01) ***
Problem →	-.03 (.01) ***		-.05 (.01) ***
Expectations →	.04 (.01) ***	-.05 (.01) ***	

Model 3 – Approaches Towards Learning, Reading Achievement and Expectations

		Next Wave	
Previous Wave	Reading	ATL	Expectations
Autoregressive			
Kindergarten	.61 (.01) ***	.40 (.01) ***	.47 (.01) ***
First Grade	.71 (.01) ***	.49 (.01) ***	.46 (.01) ***
Third Grade	.82 (.01) ***	.50 (.01) ***	.47 (.01) ***
Cross-lagged Paths			
Kindergarten			
Reading →		.22 (.01) ***	.12 (.01) ***
ATL →	.15 (.01) ***		.04 (.01) **
Expectations →	.04 (.01) ***	.04 (.01) ***	
First Grade			
Reading →		.19 (.01) ***	.12 (.02) ***
ATL →	.12 (.01) ***		.09 (.01) ***
Expectations →	.03 (.01) ***	.04 (.01) ***	
Third Grade			
Reading →		.15 (.01) ***	.17 (.02) ***
ATL →	.06 (.01) ***		.08 (.01) ***
Expectations →	.04 (.01) ***	.03 (.01) **	

Model 4 – Problem Behavior, Reading Achievement and Expectations

		Next Wave	
Previous Wave	Reading	Problem	Expectations
Autoregressive			
Kindergarten	.67 (.01) ***	.50 (.01) ***	.47 (.01) ***
First Grade	.76 (.01) ***	.53 (.01) ***	.47 (.01) ***
Third Grade	.85 (.01) ***	.52 (.01) ***	.47 (.02) ***
Cross-lagged Paths			
Kindergarten			
Reading →		-.09 (.01) ***	.14 (.01) ***

Table 28, cont.

Problem →	-.08 (.01) ***		-.02 (.01) *
Expectations →	.04 (.01) ***	-.02 (.01) *	
First Grade			
Reading →		-.10 (.01) ***	.15 (.02) ***
Problem →	-.05 (.01) ***		-.05 (.01) ***
Expectations →	.04 (.01) ***	-.04 (.01) ***	
Third Grade			
Reading →		-.08 (.01) ***	.19 (.02) ***
Problem →	-.03 (.01) ***		-.04 (.01) ***
Expectations →	.04 (.01) ***	-.04 (.01) ***	

Note. All parameter estimates are standardized with standard errors in parentheses. ATL refers to approaches towards learning.
* $p < .05$; ** $p < .01$; *** $p < .001$

Table 29. Full standardized (residual) correlations from four separate cross-lagged path models from ECLS-K without controls

Model 1 – Approaches Towards Learning, Math Achievement and Expectations			
Kindergarten	Math	ATL	Expectations
Math	1		
ATL	.45 (.01) ***	1	
Expectations	.14 (.02) ***	.10 (.01) ***	1
First Grade			
Math	1		
ATL	.17 (.01) ***	1	
Expectations	.07 (.01) ***	.07 (.01) ***	1
Third Grade			
Math	1		
ATL	.13 (.01) ***	1	
Expectations	.07 (.01) ***	.06 (.01) ***	1
Fifth Grade			
Math	1		
ATL	.11 (.01) ***	1	
Expectations	.06 (.01) ***	.08 (.01) ***	1
Model 2 – Problem Behavior, Math Achievement and Expectations			
Kindergarten	Math	Problem	Expectations
Math	1		
Problem	-.14 (.01) ***	1	
Expectations	.14 (.02) ***	-.05 (.01) ***	1
First Grade			
Math	1		
Problem	-.07 (.01) ***	1	
Expectations	.07 (.01) ***	-.03 (.01) *	1
Third Grade			
Math	1		
Problem	-.08 (.01) ***	1	
Expectations	.09 (.01) ***	-.04 (.01) **	1

Table 29, cont.

Fifth Grade			
Math	1		
Problem	-.07 (.01) ***	1	
Expectations	.07 (.01) ***	-.03 (.01) *	1
Model 3 – Approaches Towards Learning, Reading Achievement and Expectations			
Kindergarten	Reading	ATL	Expectations
Reading	1		
ATL	.42 (.01) ***	1	
Expectations	.18 (.02) ***	.10 (.01) ***	1
First Grade			
Reading	1		
ATL	.25 (.01) ***	1	
Expectations	.09 (.01) ***	.06 (.01) ***	1
Third Grade			
Reading	1		
ATL	.16 (.01) ***	1	
Expectations	.09 (.02) ***	.06 (.01) ***	1
Fifth Grade			
Reading	1		
ATL	.11 (.01) ***	1	
Expectations	.05 (.01) ***	.07 (.01) ***	1
Model 4 – Problem Behavior, Reading Achievement and Expectations			
Kindergarten	Reading	Problem	Expectations
Reading	1		
Problem	-.14 (.02) ***	1	
Expectations	.17 (.02) ***	-.05 (.01) ***	1
First Grade			
Reading	1		
Problem	-.08 (.01) ***	1	
Expectations	.09 (.01) ***	-.02 (.01) *	1
Third Grade			
Reading	1		
Problem	-.07 (.01) ***	1	

Table 29, cont.

Expectations	.10 (.01) ***	-.04 (.01) **	1
Fifth Grade			
Reading	1		
Problem	-.07 (.01) ***	1	
Expectations	.06 (.01) ***	-.02 (.01) *	1

Note. Standard errors in parentheses. ATL refers to approaches towards learning.
* $p < .05$; ** $p < .01$; *** $p < .001$

Table 30. Full standardized parameter estimates of four separate cross-lagged path models from ECLS-K with controls

Model 1 – Approaches Towards Learning, Math Achievement and Expectations			
		Next Wave	
Previous Wave	Math	ATL	Expectations
Autoregressive			
Kindergarten	.63 (.01) ***	.34 (.01) ***	.42 (.01) ***
First Grade	.67 (.01) ***	.45 (.01) ***	.42 (.01) ***
Third Grade	.80 (.01) ***	.46 (.01) ***	.42 (.02) ***
Cross-lagged Paths			
Kindergarten			
Math →		.25 (.01) ***	.07 (.01) ***
ATL →	.13 (.01) ***		.04 (.01) **
Expectations →	.02 (.01) **	.02 (.01) *	
First Grade			
Math →		.17 (.01) ***	.08 (.01) ***
ATL →	.13 (.01) ***		.08 (.01) ***
Expectations →	.03 (.01) ***	.03 (.01) **	
Third Grade			
Math →		.14 (.01) ***	.15 (.02) ***
ATL →	.06 (.01) ***		.06 (.01) ***
Expectations →	.02 (.01) ***	.02 (.01)	
Model 2 – Problem Behavior, Math Achievement and Expectations			
		Next Wave	
Previous Wave	Math	Problem	Expectations
Autoregressive			
Kindergarten	.68 (.01) ***	.47 (.01) ***	.42 (.01) ***
First Grade	.72 (.01) ***	.50 (.01) ***	.43 (.01) ***
Third Grade	.82 (.01) ***	.48 (.01) ***	.43 (.02) ***
Cross-lagged Paths			
Kindergarten			
Math →		-.07 (.01) ***	.09 (.01) ***
Problem →	-.05 (.01) ***		-.01 (.01)
Expectations →	.03 (.01) ***	-.01 (.01)	
First Grade			
Math →		-.05 (.01) ***	.11 (.01) ***

Table 30, cont.

Problem →	-.03 (.01) ***		-.03 (.01) **
Expectations →	.04 (.01) ***	-.03 (.01) **	
Third Grade			
Math →		-.04 (.01) **	.17 (.01) ***
Problem →	-.03 (.01) ***		-.03 (.01) *
Expectations →	.03 (.01) ***	-.03 (.01) **	

Model 3 – Approaches Towards Learning, Reading Achievement and Expectations

		Next Wave	
Previous Wave	Reading	ATL	Expectations
Autoregressive			
Kindergarten	.57 (.01) ***	.37 (.01) ***	.42 (.01) ***
First Grade	.64 (.01) ***	.45 (.01) ***	.42 (.01) ***
Third Grade	.78 (.01) ***	.46 (.01) ***	.43 (.02) ***
Cross-lagged Paths			
Kindergarten			
Reading →		.19 (.01) ***	.08 (.01) ***
ATL →	.14 (.01) ***		.04 (.01) **
Expectations →	.02 (.01) **	.03 (.01) **	
First Grade			
Reading →		.16 (.01) ***	.08 (.01) ***
ATL →	.11 (.01) ***		.08 (.01) ***
Expectations →	.02 (.01) *	.03 (.01) **	
Third Grade			
Reading →		.13 (.01) ***	.14 (.01) ***
ATL →	.06 (.01) ***		.07 (.01) ***
Expectations →	.03 (.01) ***	.02 (.01)	

Model 4 – Problem Behavior, Reading Achievement and Expectations

		Next Wave	
Previous Wave	Reading	Problem	Expectations
Autoregressive			
Kindergarten	.61 (.01) ***	.47 (.01) ***	.42 (.01) ***
First Grade	.68 (.01) ***	.50 (.01) ***	.43 (.01) ***
Third Grade	.80 (.01) ***	.48 (.01) ***	.43 (.02) ***
Cross-lagged Paths			
Kindergarten			
Reading →		-.07 (.01) ***	.09 (.01) ***

Table 30, cont.

Problem →	-.06 (.01) ***		-.01 (.01)
Expectations →	.03 (.01) ***	-.01 (.01)	
First Grade			
Reading →		-.07 (.01) ***	.11 (.01) ***
Problem →	-.04 (.01) ***		-.03 (.01) **
Expectations →	.02 (.01) **	-.03 (.01) **	
Third Grade			
Reading →		-.05 (.01) ***	.16 (.01) ***
Problem →	-.02 (.01) **		-.03 (.01) *
Expectations →	.03 (.01) ***	-.03 (.01) *	

Note. All parameter estimates are standardized with standard errors in parentheses. ATL refers to approaches towards learning. All models control for maternal age, type of pre-school care, disability status, child gender, and child race.
* $p < .05$; ** $p < .01$; *** $p < .001$

Table 31. Full standardized (residual) correlations from four separate cross-lagged path models from ECLS-K with controls

Model 1 – Approaches Towards Learning, Math Achievement and Expectations			
Kindergarten	Math	ATL	Expectations
Math	1		
ATL	.40 (.01) ***	1	
Expectations	.08 (.01) ***	.06 (.01) ***	1
First Grade			
Math	1		
ATL	.18 (.01) ***	1	
Expectations	.06 (.01) ***	.05 (.01) ***	1
Third Grade			
Math	1		
ATL	.15 (.01) ***	1	
Expectations	.06 (.01) ***	.05 (.01) ***	1
Fifth Grade			
Math	1		
ATL	.12 (.01) ***	1	
Expectations	.06 (.01) ***	.05 (.01) ***	1
Model 2 – Problem Behavior, Math Achievement and Expectations			
Kindergarten	Math	Problem	Expectations
Math	1		
Problem	-.12 (.01) ***	1	
Expectations	.08 (.01) ***	-.02 (.01) *	1
First Grade			
Math	1		
Problem	-.06 (.01) ***	1	
Expectations	.07 (.01) ***	-.02 (.01)	1
Third Grade			
Math	1		
Problem	-.08 (.01) ***	1	
Expectations	.07 (.01) ***	-.03 (.01) *	1

Table 31, cont.

Fifth Grade			
Math	1		
Problem	-.07 (.01) ***	1	
Expectations	.07 (.01) ***	-.01 (.01)	1
Model 3 – Approaches Towards Learning, Reading Achievement and Expectations			
Kindergarten	Reading	ATL	Expectations
Reading	1		
ATL	.35 (.01) ***	1	
Expectations	.08 (.01) ***	.06 (.01) ***	1
First Grade			
Reading	1		
ATL	.25 (.01) ***	1	
Expectations	.08 (.01) ***	.05 (.01) ***	1
Third Grade			
Reading	1		
ATL	.15 (.01) ***	1	
Expectations	.07 (.01) ***	.05 (.01) ***	1
Fifth Grade			
Reading	1		
ATL	.11 (.01) ***	1	
Expectations	.05 (.01) ***	.06 (.01) ***	1
Model 4 – Problem Behavior, Reading Achievement and Expectations			
Kindergarten	Reading	Problem	Expectations
Reading	1		
Problem	-.09 (.01) ***	1	
Expectations	.08 (.01) ***	-.02 (.01) *	1
First Grade			
Reading	1		
Problem	-.07 (.01) ***	1	
Expectations	.09 (.01) ***	-.02 (.01)	1
Third Grade			
Reading	1		
Problem	-.05 (.01) ***	1	

Table 31, cont.

Expectations	.08 (.01) ***	-.03 (.01) *	1
Fifth Grade			
Reading	1		
Problem	-.07 (.01) ***	1	
Expectations	.05 (.01) ***	-.01 (.01)	1

Note. Standard errors in parentheses. ATL refers to approaches towards learning. All models control for maternal age, type of pre-school care, disability status, child gender, and child race.
* $p < .05$; ** $p < .01$; *** $p < .001$

The results are generally consistent with those found from ECLS-B except the effect sizes are somewhat larger. Nearly every parameter is significant at the $p < .05$ level. Initial levels of parental educational expectations tend to predict increases in achievement and approaches towards learning and decreases in problem behaviors. Similarly, approach towards learning predicts increases in achievement and expectations. Problem behavior, on the other hand, predicts decreases in achievement and expectations. Achievement predicts increases in expectations and approaches towards learning and decreases in problem behavior. The results are nearly identical across models that do and do not include controls. The consistency of these longitudinal associations is particularly impressive given the massive developmental changes that children undergo in this age range, the changing school environment, the highly controlled model, and the number of parameters tested. The transactional mechanisms that undergird this type of academic development appear to be highly generalizable across outcomes (math and reading achievement), child characteristics (approaches towards learning and problem behaviors) and time (preschool to fifth grade).

As a follow-up analysis, we calculated the indirect effect of child characteristics at the initial wave that were filtered through parental expectations at the second wave on child outcomes at the final wave. Additionally, we calculated a similar pathway originating from the parent, filtered through child characteristics, and ending with ultimate expectations. To provide the most conservative estimate, we calculated the indirect effects in the cross-lagged model with controls. Ultimately, we found very small (β's roughly .005), but statistically significant indirect effects (p's < .01) for every combination of variables except for parental expectations through child problem behaviors. These indirect effects indicated that children shape their own development by influencing their parents' expectations of them, and parents shape their perceptions of their children via their effects on child behavior.

Does the transactional model differ across sociodemographic groups?

We fit a series of multiple group structural equation models to test whether parameters differed by gender, socioeconomic status quintile, and minority status. We initially fit models without cross-group constraints, and then progressively constrained autoregressive paths, and then cross-paths to equality. We tested this using a nested Satorra-Bentler scaled χ^2 difference test (Satorra, 2000), to test whether these constraints produced significant misfit. However, because the χ^2 difference test is sensitive to sample size, we also examined the CFI.

Full results can be found in Table 32. In general, we found very little evidence that the transactional pathways differed as a function of sociodemographics. For child

gender, a significant increase in model misfit was not observed for any of the four models. For minority status, a significant increase in model misfit was observed as calculated by the χ^2 difference test. However, the largest decrement in fit as assessed by CFI was only .001. Examining the group differences did not lead to any interpretable differences. Therefore, we view this as evidence that the transactional processes do not differ by student race/ethnicity. Finally, similar results were found for socioeconomic status. The χ^2 difference test indicated a significant increase in model misfit, but the largest decrement in model fit as assessed by CFI was only .007. Again, no interpretable differences were found between socioeconomic groups. These results indicate that the identified transactional processes are generalizable across most common demographic groups.

Table 32. Model Fit Comparison By Demographic Moderators

Model 1 - Approaches Towards Learning, Math Achievement and Expectations						
Gender	CFI	χ^2	*df*	True $\Delta \chi^2$	Δ *df*	*p*
Free Structure	.971	1788.929	54			
Constrain Structure	.971	1805.530	81	30.637	27	0.286
Minority Status						
Free Structure	.972	1756.007	54			
Constrain Structure	.971	1806.503	81	101.171	27	< .000
Free Autoregressive	.972	1774.417	72	37.074	18	0.005
Free Achievement	.971	1775.240	60	13.344	6	0.038
Socioeconomic Status						
Free Structure	.963	1984.381	135			
Constrain Structure	.958	2341.835	243	353.892	108	<.000
Free Autoregressive	.962	2122.167	207	114.371	72	0.001
Free Achievement	.963	2029.828	159	38.635	24	0.03

Model 2 - Problem Behavior, Math Achievement and Expectations						
Gender	CFI	χ^2	DF	True $\Delta \chi^2$	Δ DF	*p*
Free Structure	.959	2390.066	54			
Constrain Structure	.960	2355.636	81	37.573	27	0.085
Minority Status						
Free Structure	.962	2242.463	54			
Constrain Structure	.962	2281.435	81	102.72	27	<.000
Free Autoregressive	.962	2266.533	72	39.55	18	0.002
Free Achievement	.962	2263.079	60	11.137	6	0.084
Socioeconomic Status						
Free Structure	.948	2637.157	135			
Constrain Structure	.945	2879.215	243	312.762	108	<.000
Free Autoregressive	.949	2678.439	207	81.994	72	0.197

Model 3 - Approaches Towards Learning, Reading Achievement and Expectations						
Gender	CFI	χ^2	DF	True $\Delta \chi^2$	Δ DF	*p*

Table 32, cont.

Free Structure	.973	1451.921	54			
Constrain Structure	.973	1474.012	81	25.101	27	0.569
Minority Status						
Free Structure	.973	1573.492	54			
Constrain Structure	.973	1602.826	81	105.937	27	<.000
Free Autoregressive	.973	1580.904	72	29.23	18	0.046
Free Achievement	.973	1585.742	60	11.151	6	0.084
Socioeconomic Status						
Free Structure	.965	1703.936	135			
Constrain Structure	.959	2045.543	243	329.599	108	<.000
Free Autoregressive	.963	1844.471	207	113.75	72	0.001
Free Achievement	.964	1747.247	159	35.176	24	0.066

Model 4 - Problem Behavior, Reading Achievement and Expectations

Gender	CFI	χ^2	DF	True $\Delta \chi^2$	Δ DF	*p*
Free Structure	.962	1974.330	54			
Constrain Structure	.963	1952.946	81	33.614	27	0.178
Minority Status						
Free Structure	.964	2045.500	54			
Constrain Structure	.965	2033.576	81	107.649	27	<.000
Free Autoregressive	.965	2024.549	72	23.575	18	0.169
Socioeconomic Status						
Free Structure	.950	2249.324	135			
Constrain Structure	.947	2481.497	243	292.017	108	<.000
Free Autoregressive	.951	2276.095	207	68.726	72	0.588

Note. CFI stands for the comparative fit index. *df* stands for degrees of freedom. All models are compared to the model with free structure. Parameter estimates were freed in a sequential manner.

What proximal processes allow expectations to influence child academic success?

Due to the limited longitudinal data availability (measures were changed, dropped, or otherwise unavailable at other waves), we chose to examine whether parent involvement and stimulation of cognitive development statistically mediated the within-wave association between expectations and academic achievement and behavior at the first and third grade waves. We fit a dual mediation model in which the outcome was predicted by the two mediators and parental expectations. Additionally, expectations acted as a predictor of the mediators, and the mediators were allowed to correlate. This is a just identified model, and therefore fits the data exactly.

Briefly, a very small portion of the within-wave association was accounted for by parental involvement and stimulation of cognitive development. At the first grade wave, parental involvement mediated a small percentage of the total effect of expectations for math (9.8%), reading (7.7%), approaches towards learning (10.2%), and problem behavior (16.7%). At the third grade wave, parental involvement again mediated a small percentage of the total effect of expectations for math (5.5%), reading (5.6%), approaches towards learning (7.1%), and problem behavior (6.0%). Parental stimulation of cognitive development did not mediate any additional variance at either wave.

DISCUSSION

Consistent with socialization theories that emphasize parents as active transmitters of educational beliefs, we found substantial shared environmental influences

on educational expectations that were coupled with educationally-relevant behaviors and actual academic achievement. Parents who believe their children will go farther in school tend to have children who perform better academically both in terms of objective achievement and educationally relevant behaviors. Additionally, we find that parental educational expectations are themselves influenced by children's genetically influenced characteristics. This means that children are also transmitters of academic beliefs and can evoke changes in parental expectations. Our results established a complex, reciprocal pattern between child academic behaviors, child cognitive development and parental educational expectations. Our results indicate that even before entry into formal schooling, child abilities, general tendencies of academic behavior and environmental support are mutually dependent. Although discussed by a number of theoretical frameworks of child educational achievement, these reciprocal developmental mechanisms have been largely neglected in previous empirical research.

The causal ambiguity of simple correlations between parenting practices and child outcomes has been known for over 40 years. Bell (1968), for example, concluded that "the effect of children on parents can no longer be dismissed as only a logical but implausible alternative explanation of a correlation" (p. 81). Similarly, more than 30 years ago, Plomin et al. (1977) and Scarr and McCartney (1983) provided a motivating developmental theory that incorporates processes that link a child's genotype with the types of environments that they passively receive from their parents, evoke from their surroundings, and actively seek out in accord with their genetic predispositions. Indeed, even some elaborations of the E-V model have allowed for the possibility that "parents'

and children's beliefs are likely to influence each other reciprocally" (Jacobs & Eccles, 2000, p. 416). Nevertheless, rigorous empirical research to test such hypotheses remains rare. (Perhaps implicitly) guided by unidirectional transmissive thinking, much of the empirical work on parental educational expectations takes the approach of measuring parental expectations before child variables - an approach incapable of testing for child-to-parent effects. Why transactional frameworks have not been better integrated into empirical studies of academic development is unclear. J. Richard Udry (2003), in a characteristic to-the-point style, provided his opinion. He argued:

> Most social science theories assume parent-to-child effects as the basic causal sequence because they do not believe that children have inherent attributes. If children do not have inherent attributes, then there is no starting point in the child. It is tabula rasa all over again….If you believe that individuals differ from one another from birth because of inherent attributes, then no assumption of parent-to-child as the starting point makes any sense. Longitudinal designs will not solve the problem. Nor will starting your investigations at younger and younger ages. (p. 49)

Our aim for the current project was to provide an empirical counterpoint to the overwhelming tendency to test unidirectional parent-to-child models of parental educational expectations and child academic development. We believe that empirical research should strive to match the complexity of the theory and child development.

"Heritable" Environments and Gene-Environment Correlation

We found that a significant portion of variance in parental beliefs about the educational future of their children was associated with child genotypic differences. Because parental beliefs are traditionally conceptualized as environmental contexts, this result is indicative of active or evocative gene-environment correlation, whereby parents form their expectations on the basis of genetically influenced characteristics of their children. In line with previous research indicating that cognitive ability influences the type and quality of parental interaction that children receive (Lugo-Gil & Tamis-LeMonda, 2008; Tucker-Drob & Harden, 2012a), our results indicate that parents are sensitive to their children's math and reading achievement and adjust their expectations over time accordingly. We also found evidence that parental beliefs are formed, in part, based on a child's general tendency for behaviors that facilitate or hinder task-focused academic learning. These genetically influenced child characteristics predicted later parental educational expectations, allowing the genetically influenced behaviors to get "out of the skin" to influence environmental experience, even before children entered formal schooling. Interestingly, while our behavior genetic models indicated shared environmental mediation of the associations between achievement and educational expectations, approaches towards learning and educational expectations consistently reflected shared genetic, rather than environmental, influences. Thus, our results suggest that general patterns of behavior, as opposed to achievement test scores, may be more

robust mechanisms of gene-environment correlation with respect to parental expectations in early childhood.

Despite the significant amount of variance in parental educational expectations associated with child genotype, shared environmental effects accounted for the majority of the variance in expectations. Conceptually, these influences represent child-invariant, family-level influences on parenting. The mechanisms leading to these large shared environmental influences on educational expectations, however, are not entirely known. We found that only about one quarter of the shared environmental effect on expectations could be attributed to sociodemographics, such as socioeconomic status and race/ethnicity. The remaining variance could be due to parenting values, other family-level cultural or environmental influences, or genetically influenced traits of the parent (e.g., Avinun, Ebstein, & Knafo, 2012; Bakermans-Kranenburg & Van IJzendoorn, 2008). If parents possess genetically influenced traits that affect their parenting and these traits are passed on to their children, this represents a type of passive gene-environment correlation. As it operates to make children living in the same household more similar to one another, passive gene-environmental correlation would act as a shared environmental influence even though the developmental process is partially genetic. A more complex design would be necessary to evaluate this possibility, for instance a children-of-twins design. It is likely that the shared environmental factor represents several of these possibilities. The nonshared environment, representing differential within-family treatment, was very small. This indicates that, apart from differences associated with

genotypic differences between children, parents form very similar expectations for their children.

Incorporating gene-environment correlation into academic socialization models has important implications for developmental theory. One of the most widely replicated findings in all of behavior genetics is that the heritability of cognitively-relevant outcomes increases with age (Briley & Tucker-Drob, 2013; Haworth et al., 2010). Explanations for this finding rely primarily on an understanding of active gene-environment correlation whereby children increasingly select environments that are congruent with their genetic predispositions as they age (Dickens & Flynn, 2001; Scarr & McCartney, 1983). Applied to the current results, preliminary evidence for this process can be found in the longitudinal models. In general, the effect sizes are larger in the ECLS-K dataset (of elementary and middle school development) compared to ECLS-B (of preschool and kindergarten development), which is consistent with the hypothesis that as children age, they exert an increasing influence on their own achievement and the type of received parental support. As this feedback process depends, in some small part, on genetically influenced characteristics, it is likely that the heritability of expectations would increase along with other cognitive variables. Crucially, this implies that the heritability coefficient is not a deterministic value that limits socialization effects, but rather, heritability depends on socialization processes to guide child development (see Bronfenbrenner & Ceci, 1994; Tucker-Drob et al., 2013).

Both Children and Parents as Drivers of Academic Development

The present results indicate that children are important drivers of the climate of their academic development. We have focused on child driven effects for the majority of the article because this pathway is often overlooked in empirical studies. It is also important to emphasize that the current results implicate parents as strong drivers of academic development. We found a significant shared environmental correlation between parental educational expectations and child math and reading achievement both within waves and across waves of ECLS-B. Transmitted beliefs, values and perceptions of competence are likely mechanisms for this shared environmental correlation. Therefore, we would argue that our results are largely consistent with previous work on educational expectations. An open empirical question, however, is to what extent the shared environmental effects detected reflect patterns of passive gene-environment correlation (described above).

The longitudinal cross-lagged path models identified bi-directional interactions between parents and children, but it is somewhat unclear whether genetic or environmental mechanisms mediate these associations. For example, parent educational expectations may be associated with later approaches towards learning because of an underlying genetic pathway, consistent with the idea that parents are responding to enduring patterns of behavior in their children. Several models of cognitive development (e.g., Bouchard, 1997; Dickens & Flynn, 2001; Hayes, 1962; Scarr, 1997) speculate that only environments that are experienced as a result of this type of gene-environment

correlation are likely to have an appreciable influence on development as these environments will be recurrently experienced with age. Unfortunately, we were unable to decompose the cross-lagged pathways into variance components due to genetic, shared environmental, and nonshared environmental factors. A much larger sample size of twins would be required to detect these effects at this age.

Locating Causal Effects in Development

Our results indicate that the dynamic processes between the student and their environment begin to shape academic trajectories even before the entry into schooling. It is therefore possible that correlations found in older students may largely reflect the accumulated effects of processes that are initiated very early in childhood. Our cross-lagged path models provide support for the concept of a "developmental cascade" (e.g., Bornstein, Hahn, & Wolke, 2013). Another recent example of such a process comes from Bornstein, Hahn, and Suwalsky (2013), who found that exploratory ability of 5 month old infants was associated with academic achievement in adolescence through intermediate associations with intellectual development. One explanation for this result is that children vary in their ability or tendency to explore their world and actively seek out or evoke environmental experiences (Raine, Reynolds, Venables, & Mednick, 2002). If individual differences in motor ability are even slightly genetically influenced, then such a developmental cascade could result in variation in ability becoming increasingly tied to genotypic differences. We found that very early indicators of a child's achievement and

behavioral tendencies predict change in their academic trajectory, as do very early parental influences.

In ECLS-B (ages 4 and 5), we found that the effect of early parental expectations on later child achievement was more than double the effect of early achievement on later parenting. In ECLS-K (grades K through 5), however, the child-to-parent effect grew to more than four times *larger* than the parent-to-child effect for academic achievement. This likely reflects a response to growing divergence of student academic trajectories. Importantly, the parent-to-child parameter was nearly identical across the separate ECLS-B and ECLS-K datasets, but differences in the child-to-parent effect account for this striking difference. Parents likely generate stable expectations for their children at an early age, but over time parents dynamically adjust their expectations such that levels of expectations become increasingly child based. Further, the influence of early child academic behaviors on later achievement was more than twice as large in ECLS-K as in ECLS-B. Again, growing divergence in academic trajectories emerges from magnified differences in early patterns of behavior. Stratification of achievement may result from the dynamic interaction between child predispositions for learning, child ability, and their educational environment. Focusing attention towards these types of early transactions between children and their environments may prove beneficial for research that aims to foster upward trajectories of academic achievement.

Strengths and Limitations

This study has a number of strengths that support the conclusions being drawn. We applied both behavior genetic models and cross-lagged path models to high quality, population representative, longitudinal data of educationally relevant outcomes. The findings of behavioral genetic studies of education are rarely integrated within socialization frameworks of child development. We view these models and methods to be highly complementary and provide unique information about child development. Moreover, as all modeling approaches are limited by their unique sets of assumptions, our inferences are strengthened by having been conceptually replicated across behavioral genetic and longitudinal approaches.

Several limitations are of note. We were unable to evaluate whether the gender of the parent or teacher matter for the report of academic behavior. The vast majority of respondents were female (>95%) in ECLS-B. In ECLS-K, the respondent gender was suppressed except at the initial wave when the teachers were almost entirely female (98%). Understanding how the socialization process unfolds in relation to male and female parents is an important future direction for research.

Our analyses of child academic behaviors were somewhat limited in the ECLS-B dataset because we relied on parent report of behavior. Teacher report of academic behavior was available for some participants, but children in formal pre-school differ systematically from the general population (see Tucker-Drob, 2012). The behavioral genetic results were largely similar across parent and teacher report indicating that shared method variance between parent report of behavior and expectations does not fully

explain the results. Additionally, the highly consistent results from ELCS-K were fully based on teacher report of academic behavior.

We were unable to fully explain the child genetic influences on educational expectations. There is much left to be explained both in terms of the gene-environment interplay in the formation of expectations for parents and children, as well as the mechanism by which academic behaviors, academic achievement, and expectations are prospectively related. In particular, the cross-lagged path models presented here add clarity to the directionality of effects, but it is unclear whether genetic or environmental mechanisms link the outcomes across time. A much larger twin sample would be necessary to fully integrate the behavior genetic and the longitudinal cross-lagged approach. Similarly, we found that early parent and child characteristics could predict later outcomes, but residual associations remained suggesting influential unmeasured factors. Unraveling this association across time is an important avenue for future research. Importantly, the finding that child genetic differences influence expectations does not invalidate the "importance" of parents; it displays the developmental process more accurately. Parents still play an active role, for example, by being receptive and open to forming educational beliefs on the basis of their child's preferences.

We were able to evaluate a large span of child development by combining the ECLS-B and ECLS-K datasets. We tracked children across the transition into kindergarten through fifth grade and nearly every parameter was in the expected direction with the majority statistically significant. The results are somewhat limited due to the fact

that ECLS-B and ECLS-K are separate datasets. Therefore, we cannot draw longitudinal inferences across datasets.

One may wonder whether the effects uncovered in the present study are too small to have a substantial impact on child development. For example, genes only accounted for roughly 20% of the variance in parental educational expectations, and cross-lagged paths were typically small. However, these effects may have more practical importance than might be expected at face. Because the large majority of individuals pass through the education system, even very small benefits are likely to pay large dividends to society. The promise of a transactional model is that reciprocal feedback loops can be constructed to facilitate compounding benefits with development. In this sense, small effects over one year intervals may translate to large effects over the entirety of development. Further, expectations represent only one environment that may be selected on the basis of genotype. Evaluating gene-environment interplay for other academically relevant beliefs and values may add to the transactional model presented here.

Finally, the use of a single indicator to assess the child's academic environment has important tradeoffs. On the one hand, educational expectations are a distal factor that likely influences many more proximal behaviors, beliefs, and values that shape child development. Although this may limit mechanistic interpretations, it is clear from the lack of variance mediated by specific proximal behaviors (i.e., involvement and stimulation) that the use of such a broad variable indexes a wide array parental inputs. On the other hand, parental educational expectations may reflect a more narrow measure than broad sociodemographic factors, such as socioeconomic status and race/ethnicity which

tended to attenuate the influence of expectations. In this case, knowing that parental expectations predicts achievement in a heterogeneous population provides additional insight into the social context of child development than attributing this effect to broad group differences.

Conclusion

The present study made use of behavior genetic and longitudinal methodology to address whether children actively evoke changes in parental beliefs and influence their developmental environment. We tested these plausible, but previously unexplored, connections between children and parents and found strong evidence that child-to-parent effects do influence educational expectations. Our results are consistent with a fully transactional model between child academic behaviors, child academic development and parental educational expectations that shapes the educational trajectories of children. Even before entry into formal schooling, children influence their educational environments.

Chapter 4: Geographic Variation in Personality is Associated with Fertility across the United States

Geographical regions across the United States differ widely in their fertility norms—how many children individuals have, when they have children, what control over fertility is seen as morally permissible, and the family contexts in which fertility takes place. Conventional explanations center on interactions between regional institutional, political, economic, and religious forces and individual behavior (D'Addio & D'Ercole, 2005; Glass & Levchak, 2014; Lesthaeghe, 2014; Lesthaeghe & Neidert, 2006; Morgan, 1996). A nascent body of research indicates that personality—contextually and developmentally stable patterns of thinking, feeling, and behaving (John, Naumann, & Soto, 2008)—is predictive of fertility outcomes at the individual-level in both human (Berg, Rotkirch, Väisänen, & Jokela, 2013; Courtiol, Pettay, Jokela, Rotkirch, & Lummaa, 2012; Hutteman, Bleidorn, Penke, & Denissen, 2013; Jokela, 2012; Jokela, Kivimäki, Elovainio, & Keltikangas-Järvinen, 2009) and non-human populations (Aplin et al., 2013; Réale, Martin, Coltman, Poissant, & Festa-Bianchet, 2009; Seyfarth, Silk, & Cheney, 2012). However, no research has yet considered geographic variation in personality as an incremental predictor of regional fertility. Using a sample of 890,253 U.S. residents, here we show that states with relatively high average levels of extraversion, agreeableness, and conscientiousness and lower levels of neuroticism and openness tend to display more traditional fertility (i.e., higher fertility, earlier fertility, and more structured practices). Fertility has a tremendous impact on the well-being of society through effects on national growth, economic stability, and population aging (Bloom, Canning, Fink, & Finlay, 2010; Demeny, 2003; Harper, 2014; Morgan, 2010). Moreover, reproductive health policies are consistently hotly debated issues in science, politics, and culture (Mills, Rindfuss, McDonald, & te Velde, 2011). These findings point

to an entirely new kind of predictor of reproductive rates with important implications for population projections, cultural divides surrounding fertility practices, and evolutionary models of fluctuating selection pressures resulting from ecological social niches.

To more accurately model geographic variation in fertility, an outcome with importance for economics, population health, and evolutionary theory, we consider regional variation in personality. In the United States, some regions have more extraverted (e.g., outgoing vs. timid), agreeable (e.g., warm vs. confrontational), conscientious (e.g., disciplined vs. accidental), neurotic (e.g., anxious/depressed vs. emotionally stable), and open (e.g., creative/intellectual vs. conventional) individuals than do other regions. This geographic variation may result from causal processes, in which social and structural characteristics of regions affect personality development, and from selective processes, in which individuals with specific personality traits systematically migrate to and away from different regions (Rentfrow, Gosling, & Potter, 2008). Geographic variation in personality is predictive of such diverse outcomes as presidential voting patterns (Rentfrow, Jost, Gosling, & Potter, 2009), entrepreneurial activity (Obschonka, Schmitt-Rodermund, Silbereisen, Gosling, & Potter, 2013), and several other political, religious, economic, sociological, and health indicators (Rentfrow et al., 2013). The associations between social phenomena and concentrations of personality may result from both top-down influences of social institutions on psychological development (e.g., living in an active artistic community affects levels of openness) and bottom-up influences of personality on the creation of social structures and outcomes (e.g., concentrations of highly open individuals generate artistic social settings). Therefore, links between regional personality and fertility may emerge from shared ecological-level influences (e.g., influential religious institutions increase both levels of fertility and agreeableness) and from the aggregation of individual-level

personality effects on fertility (e.g., agreeable individuals tend to form certain family types, and this effect sums across many individuals living in a region). To date, no study has examined such links.

We constructed measures of state-level extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience based on geographically coded data from 890,253 U.S. participants who responded to the Big Five Inventory (John et al., 2008). The reliability and validity of this instrument across the U.S. has been well-documented in previous research (see Methods; Rentfrow et al., 2008; Rentfrow et al., 2013).

Based on data from the National Vital Statistics System, we primarily focus on the total fertility rate (i.e., the average number of children that would be born to a woman if she experienced the age-specific fertility rates that prevailed in a given period through her lifetime). This rate has the most direct impact on the global population through cohort replacement, and subreplacement fertility (i.e., total fertility rates below 2.1) could restructure the age distribution of the population causing economic instability (Bloom et al., 2010; Demeny, 2003; Harper, 2014; Morgan, 2010). In areas with low fertility, cohorts are not replaced causing population aging, economic burden associated with health care, and loss of productivity.

We also included as outcomes other features of the fertility schedule (highlighted in Figure 17; Schmertmann, 2003) and markers of fertility-relevant behaviors identified by previous research as central to heterogeneous regional fertility (Lesthaeghe & Neidert, 2006). These included age at first birth, age at first marriage, percent never married, percent of marriages that ended in divorce, the percentage of cohabiting households, non-marital fertility rate, percent unintended pregnancy, abortion rate, and family planning expenditures per woman in need of contraceptives.

We selected a comprehensive set of conventional sociocultural predictors to include in our analyses. Regional differences in fertility are known to be associated with demographic, political, and religious characteristics (Lesthaeghe & Neidert, 2006). Therefore, we included a number of state-level predictors: median household income, percent African American, percent Hispanic, percent female, percent that has obtained a college degree, the percent that lives in an urban area, the percent that report that religion is very important to them, and the percent that voted for Obama in the 2008 election.

As a preliminary step, we compared the predictive power of the personality variables to that of the conventional sociocultural predictors. We find that the five personality variables explain 52% of the between-state variation in total fertility. That this percentage is nearly as large as that explained by conventional predictors ($R^2 = .57$) is particularly striking, given that personality has never before been implicated in geographical variation in fertility. Moreover, a regression that includes both personality and conventional predictors explains 74% of the between-state variation, indicating that both personality and conventional predictors account for variation in fertility uniquely of one another. In the remainder of this letter, we report incremental associations between personality and fertility outcomes. As described in the Methods, we accomplished this by residualizing the main study variables for all conventional predictors. Thus, the effect sizes reported below can be considered conservative, because regional personality may have indirect effects on fertility through values or policy (Rentfrow et al., 2009).

Table 33 reports correlations between personality and fertility outcomes, adjusted for all conventional predictor variables. Agreeableness, conscientiousness, and openness to experience explained significant portions of variance in total fertility with large effect sizes ($r \sim |.50|$). Put differently, these effects indicate that each 1 SD unit difference in regional personality translates to a difference of approximately .07 children per woman in

a state. Total fertility across the United States ranged from 1.63 to 2.45 (SD = .17) in 2010, meaning that the difference associated with 1 SD change in personality amounts to 9% of the observed range. The low end of this range indicates subreplacement fertility and substantial population aging, but at the high end, population growth is expected. The current results imply that a change from low levels of regional personality (i.e., 2 SD below the mean) to high levels (i.e., 2 SD above the mean) would lead to a difference of .28 children per woman. Such a change would raise 78% of the 41 states with subreplacement fertility above this level. Trending correlations ($p < .10$) with total fertility were also observed for extraversion and neuroticism with moderate effect sizes ($r \sim |.25|$). States with high total fertility were marked by high extraversion, agreeableness, and conscientiousness, and low neuroticism and openness.

Table 33. Correlations between personality and fertility outcomes

	E	A	C	N	O
Total fertility rate	.24	**.51**	**.42**	-.25	**-.53**
Initiation Age	-.17	.07	-.05	-.07	-.04
Peak Fertility Age	.15	.09	.21	.01	.02
Stopping After Peak Fertility	**.36**	.09	**.36**	-.03	-.22
Age at first birth	.05	.12	.18	**.38**	-.04
Age at first marriage	-.20	-.09	-.11	**.37**	.17
Percent never married	-.16	-.07	-.18	.04	**.32**
Percent divorce	.02	-.01	-.02	-.05	-.01
Percent cohabit	-.20	-.21	-.25	**.31**	.21
Non-marital fertility rate	**-.32**	-.23	**-.38**	.00	.20
Percent unintended pregnancy	-.25	.04	-.15	.12	.01
Abortion rate	.04	.07	.14	**.36**	-.17
Family planning expenditures	-.23	.04	-.17	-.04	.19

Note. E = Extraversion. A = Agreeableness. C = Conscientiousness. N = Neuroticism. O = Openness to experience. All variables adjusted for sociodemographic characteristics and value controls. Parameters printed in bold are significant at $p < .05$.

Turning toward fertility-relevant behaviors, higher state-level neuroticism was associated with later age at first birth and marriage and higher rates of cohabitation and abortion. States with higher openness tended to have larger never married populations. These moderate to large associations ($r > .30$) indicate that state-level neuroticism and openness tend to be associated with markers of non-traditional fertility, particularly in reference to delayed family formation. Extraversion and conscientiousness predicted greater stopping behavior and lower non-marital fertility rates ($r \sim |.35|$). Non-trivial negative correlations ($r < -.22$) were found between extraversion and unintended pregnancy, between agreeableness and non-marital fertility, and between conscientiousness and cohabitation. These weak to moderate associations point toward extraversion, agreeableness, and conscientiousness as markers of traditional fertility, particularly in reference to family structures where fertility occurs. Figure 16 presents scatterplots for each personality trait and a major correlate.

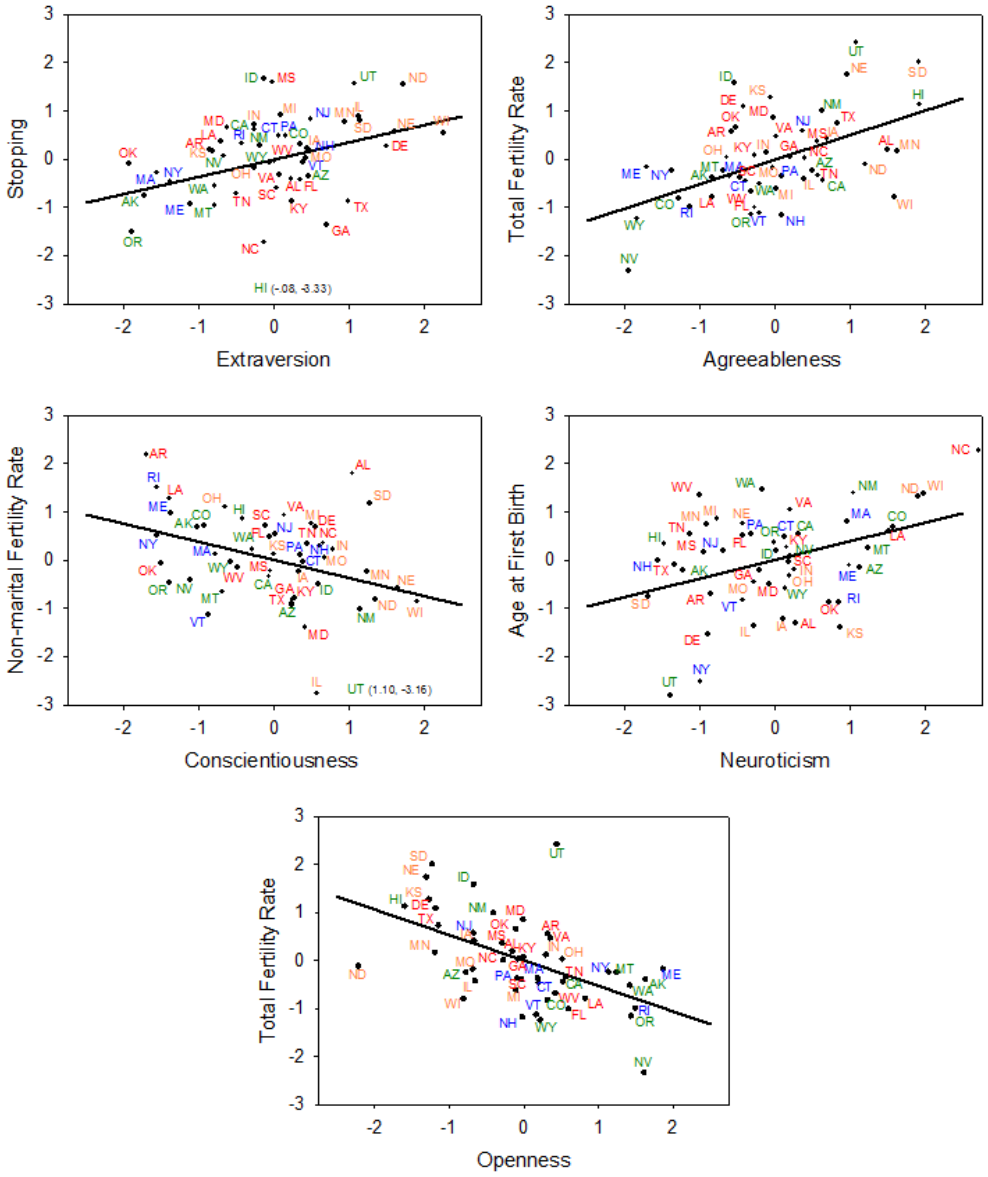

Figure 16. Scatterplots of fertility outcomes by personality. All variables are adjusted for sociodemographic characteristics and value controls and plotted as standardized residuals. Data points are labeled with associated state abbreviations and colored based on Census regional designations (Northeast = blue; Midwest = orange; South = red; West = green). Hawaii is a small outlier for stopping (z = -3.33), and Utah is a small outlier for non-marital fertility (z = -3.15).

We tested whether the Big Five personality factors uniquely or commonly explained variance in fertility. Table 35 presents standardized regression parameters from this analysis, and the pattern of results largely matches Table 33. Common variance among personality accounted for roughly half of the effect on the total fertility rate and non-marital fertility rate. The remaining outcomes primarily reflected unique associations with individual personality factors.

Our results are inconsistent with the hypothesis that regional-level personality-fertility associations simply and exclusively represent the aggregation of individual-level effects. For instance, individual-level studies find that conscientiousness is associated with *lower* fertility (Jokela, 2012). We find the opposite result at the region-level; conscientiousness tends to predict higher total fertility rates. Further, individual-level studies find that extraversion is associated with increased likelihood of unintended pregnancies (Berg et al., 2013). We find state-level extraversion is associated with lower unintended pregnancies. Of course, mechanisms leading to individual-level and region-level phenomena do not necessarily depend on one another.

Our results are more consistent with the hypothesis that regional differences in personality influence regional policies and social norms that in turn affect individual-level fertility outcomes. In other words, states differ in terms of the social climate of fertility beliefs, public policy, and other sociodemographic predictors of fertility (D'Addio & D'Ercole, 2005; Glass & Levchak, 2014; Lesthaeghe, 2014; Lesthaeghe & Neidert, 2006; Morgan, 1996). Individuals tend to create these institutions and general social contexts partially on the basis of individual differences in personality (Rentfrow et al., 2008). These societal institutions may exert top-down influences on individual-level fertility outcomes.

Fertility differentials may also influence personality concentrations. In addition to personality predicting subsequent fertility, the experience of parenthood results in personality change (Jokela et al., 2009). Thus, having a child or even living in a region that emphasizes childbearing may change personality levels, and therefore create a link between fertility and regional personality (Bleidorn et al., 2013).

Finally, it is of note that personality and fertility outcomes are partially heritable (Briley & Tucker-Drob, 2014; Harden, 2014), and therefore regional concentrations of these phenotypes might emerge from differential patterns of migration that persist across generations (i.e., founder effects). Historically, the spread of sociocultural influences relevant to fertility also followed migration flows (Woodard, 2011). If levels of personality in the population influence the creation of norms, genetic and sociocultural transmission could combine to produce regionally distinct fertility practices linked to personality. This implies that evolutionary selection pressures may vary across geographical space in response to culturally created ecological niches in modern societies. This effect may operate independently or jointly with variation associated with environmental pressures or resources.

Reproductive behavior shapes the future of society. Economic and public policy decisions often rely on demographic forecasts of population growth, development, and aging based on known determinants of fertility. For the first time, we add regional personality as a strong and entirely independent predictor of fertility. Our implementation of an extensive set of known correlates of fertility ensured that the detected personality-fertility associations were completely novel. Future work will be necessary to disentangle the specific mechanisms underlying these new links with fertility. Theoretically, the present results highlight the dynamic interplay between socially constructed ecological niches and fertility behavior, which has implications for understanding the unfolding of

evolution across chronological time and geographic space. Our findings open new avenues for research on the mechanisms of persistent geographical heterogeneity in fertility and for modeling population growth and geographical dispersion.

Methods

Regional Estimates of Personality

We obtained regional estimates of personality from a very large scale, online study (Gosling, Vazire, Srivastava, & John, 2004). Self-reports on the Big Five Inventory (John et al., 2008) were obtained from 890,253 individuals in the United States. No statistical method was used to determine sample size, as larger samples provide more stable estimates of state-level personality. All participants provided informed consent approved by the Institutional Review Board at the University of Texas at Austin. Responses were classified based on reported state of residence. Numerous measures have been taken to ensure the validity, representativeness, and reliability of the data. These procedures are described in several publications (Rentfrow et al., 2008; Rentfrow et al., 2009; Obschonka et al., 2013; Rentfrow et al., 2013). We controlled for the influence of response sets, such as acquiescence (i.e., yea-saying) and extreme responding (i.e., preferential use of polar response options; John et al., 2008). Additionally, we controlled for the individual-level influence of age, age^2, gender, and an age-×-gender interaction so that our analyses would not be confounded by demographic differences of the sample. From this individual-level data, we calculated state-level aggregates for the Big Five.

In addition, we created separate measures based on segments of each state's population. Associations between personality and fertility differ across gender at the individual-level (Jokela, 2012). Regional personality levels of males and females may have differential associations with fertility because of gendered divisions of labor and

childrearing (McDonald, 2000). For example, regional female conscientiousness may have associations with fertility independent of male conscientiousness due to mechanisms linked to gender roles concerning child care. Similarly, fertility rates follow a strong age pattern (see Figure 17). Regional personality levels of younger and older individuals may have differential associations with fertility because the older population typically has greater control over policy and institutions (Ingraham, 2014), but the younger population is responsible for the majority of actual births (Martin et al., 2012). For example, regional openness of the younger population may have associations with fertility independent of the older population due to mechanisms linked to reproductive behavior. We explored these potential driving mechanisms of personality-fertility associations. Specifically, we created state-level personality measures for male and female individuals, the younger population (age < 30) and the older population (age > 30), and the difference between gendered (male personality - female personality) and aged personality (younger personality - older personality). Thus, we calculated a total of 7 (data conditions) × 5 (Big Five) estimates of personality for each state.

Fertility Schedule

We obtained 5-year age-specific fertility rates for each of the 50 states for the year 2010 (Martin et al., 2012). We transformed the 5-year age-specific fertility rates into 1-year age-specific fertility rates using the method designed by Schmertmann (2012). This method uses historical consistencies in fertility schedules to estimate the most likely 1-year age-specific fertility rates. From this, we fit Schmertmann's (2003) calibrated spline model to the fertility schedules to provide parameters that are intuitively meaningful. This model uses very few parameters to construct a continuous fertility function. We focus on four aspects of the fertility schedule. First, the total fertility rate represents the

average number of children that would be born to a woman if she experienced the age-specific fertility rates that prevailed in the year 2010 through her lifetime. This reflects the overall level of fertility in a given state. Second, initiation reflects the earliest age at which fertility begins. Third, peak fertility refers to the age at which fertility is highest. Fourth, stopping refers to the force of individuals controlling maximum fertility, presumably after a desired family size has been reached. Following Schmertmann's[26] recommendation, stopping is calculated as the difference between the age at which fertility would linearly fall to half from peak fertility to age 50 and the actual age at which fertility reaches half of the peak. Larger stopping values indicate a steeper decline in fertility following the peak and presumably more control of fertility. The initiation, peak, and stopping parameters describe differences in the shape of the fertility schedule. Figure 17 displays 8 example distributions highlighting each parameter.

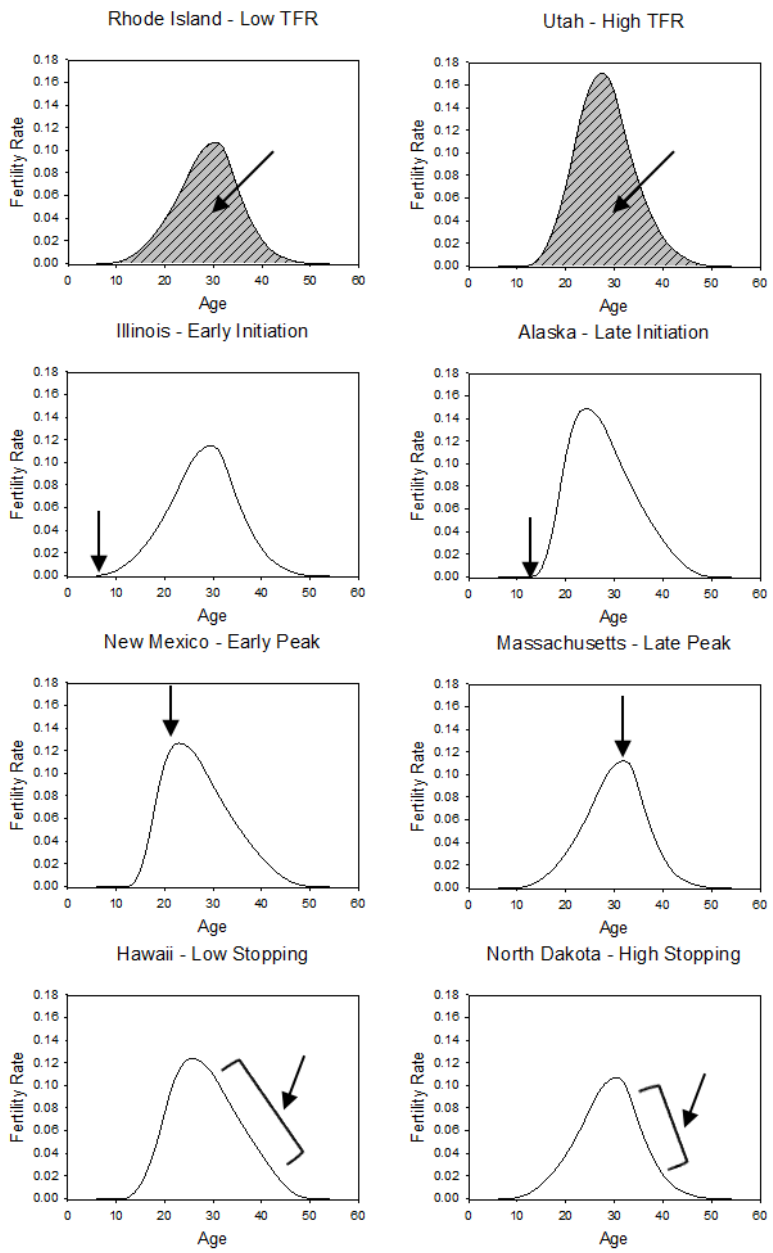

Figure 17. Example fertility schedules for eight states with lowest and highest values for total fertility, initiation, peak, and stopping. TFR = Total Fertility Rate. TFR represents the area under the curve. Initiation represents the earliest age with a non-zero fertility rate. Peak represents the age that fertility is highest. Stopping represents the force of limiting fertility after peak fertility, which is conceptually analogous to the slope of the curve after the peak. Data from National Vital Statistics System.

Fertility-Relevant Behaviors

We included several markers of behaviors previously found to be indicative of regional variation in fertility (Lesthaeghe & Neidert, 2006). This includes age at first birth, age at first marriage, percent of the population never married, percent of marriages that ended in divorce in the last year (in reference to the total married population), the

percentage of cohabiting households, non-marital fertility rate, percent unintended pregnancies, abortion rate (i.e., number of abortions per 1,000 women aged 15-44), and family planning expenditures per woman in need of contraceptives (Finer & Kost, 2011; Jones & Kooistra, 2011; Lofquist, Lugaila, O'Connell, & Feliz, 2010; Mathews & Hamilton, 2009; Sonfield & Gold, 2012). Most indicators were obtained for the year 2010 and are based on 2010 U.S. Census estimates or on the American Community Survey.

Conventional Predictors

As conventional predictors, we included state differences in median household income, percent African American population, percent Hispanic population, percent female population, percent of population that has obtained a college degree, and the percent of the population that lives in an urban area based on estimates from the 2010 U.S. Census. Based on previous evidence that regional variation in fertility is associated with values (Lesthaeghe & Neidert, 2006), we additionally included the percent that voted for Obama in the 2008 election (Federal Election Commission, 2009) and the percent that report that religion is very important to them in the Gallup (2010) State of the States poll.

Data Preparation

Table 34 presents descriptive statistics for all study variables, including measures of spatial autocorrelation using Moran's I (Moran, 1950). This statistic indicates the degree to which observations are more similar to geographic neighbors than would be expected by chance. Positive values indicate that high values tend to cluster near high values and low values tend to cluster near low values. Negative values indicate that observations are dispersed, with high values near low values and vice versa.

Nonsignificant Moran's I values indicate that observations are distributed randomly across geographic space. Interpreting and addressing spatial autocorrelation is important because it indicates that empirical observations are not truly independent of one another and may lead to faulty statistical inferences due to violations of statistical assumptions (e.g., independent and identically distributed errors; Anselin, 1988).

To calculate Moran's I, we used a queen's case contiguity weights matrix and TIGER shapefiles from the U.S. Census Bureau (2010). We used the spdep package in R to estimate Moran's I (Bivand, 2014). We allowed for geographic islands (e.g., Alaska and Hawaii). Generally, we observed positive spatial autocorrelation for nearly all of the study variables.

We were primarily interested in the association between state-level personality and fertility, holding known correlates constant. Therefore, we computed residuals from linear models in which each of the primary study variables (i.e., personality and fertility) were predicted by all of the conventional predictors. The residual variance from these models is reported in Table 34. The conventional variables accounted for approximately half of the variance of personality and fertility, on average. This indicates that sociodemographic characteristics and values are important explanatory variables for both state-level personality and fertility, but substantial unexplained variation remains. Because we were not interested in the regression parameters from these models (which may be biased due to spatial autocorrelation), we used ordinary least squares regression.

We then tested for spatial autocorrelation among the residuals. For most variables, controlling for sociodemographic characteristics and values reduced spatial autocorrelation to nonsignificance. Remaining spatial autocorrelation was observed for neuroticism, female neuroticism, peak fertility, age at first birth, percent never married, and percent divorce. To avoid spatial autocorrelation obscuring associations between

personality and fertility, we turned to spatial regression models which take into account the spatial structure of the data (LeSage & Pace, 2009). These include spatial lag models, where the spatial lag of the dependent measure is included as a predictor, spatial error models, where errors are allowed to correlate based on geographical relations, and models that combine both (i.e., SARAR). Each model is designed to control for the spatial structure of the data and provide unbiased parameter estimates. Spatially oriented regression diagnostics can detect sources of geographic error and guide selection of the appropriate model (Anselin, Bera, Floras, & Yoon, 1996). Based on these diagnostic tools, we used a spatial error model to predict female neuroticism, peak fertility, age at first birth, percent never married, and percent divorce by the sociodemographic characteristics and value controls, and we used a spatial lag model to predict neuroticism by the controls. Residuals from these models did not show spatial autocorrelation and were saved for analysis.

To summarize, most study variables were spatially autocorrelated across the United States and shared variance with sociodemographic characteristics and control variables. We used ordinary least squares regression to predict each study variable by the controls and saved the residuals for analysis. If the residuals still displayed spatial autocorrelation, we used spatial regression models to account for the spatial structure of the data. This was the case for 6 study variables. Ordinary least squares regression produced non-spatially autocorrelated residuals for the other variables. Use of standard correlation and regression techniques for the primary analyses is valid because no variable displayed spatial autocorrelation.

Table 34. Descriptive statistics of study variables

Variable	Mean (SD)	Range	Year	Moran's I	Residual Variance	Residual Moran's I
Personality						
Extraversion	0.00 (0.04)	-0.13-0.08	1999-2005	.22 *	.71	.14
Agreeableness	0.00 (0.06)	-0.15-0.11	1999-2005	.25 **	.75	.12
Conscientiousness	0.00 (0.05)	-0.11-0.09	1999-2005	.34 ***	.66	.14
Neuroticism	0.00 (0.05)	-0.08-0.14	1999-2005	.52 ***	.43	.19 *
Openness to experience	-0.02 (0.06)	-0.17-0.09	1999-2005	.35 ***	.47	.15
Male Personality						
Extraversion	0.09 (0.03)	0.00-0.17	1999-2005	.24 **	.69	.16
Agreeableness	0.47 (0.03)	0.39-0.52	1999-2005	.09	.77	.03
Conscientiousness	0.32 (0.03)	0.26-0.38	1999-2005	.40 ***	.50	.00
Neuroticism	-0.29 (0.03)	-0.34--0.22	1999-2005	.50 ***	.53	.15
Openness to experience	0.43 (0.03)	0.37-0.51	1999-2005	.29 **	.47	.00
Female Personality						
Extraversion	0.21 (0.02)	0.14-0.26	1999-2005	.20 *	.77	.10
Agreeableness	0.58 (0.03)	0.50-0.64	1999-2005	.31 ***	.72	.14
Conscientiousness	0.41 (0.03)	0.34-0.46	1999-2005	.31 ***	.62	.13
Neuroticism	0.02 (0.03)	-0.04-0.12	1999-2005	.51 ***	.41	.18 *
Openness to experience	0.36 (0.04)	0.27-0.42	1999-2005	.36 ***	.52	.16
Less than 30 Personality						
Extraversion	-0.01 (0.04)	-0.14-0.09	1999-2005	.23 **	.63	.12
Agreeableness	0.01 (0.07)	-0.16-0.15	1999-2005	.24 **	.75	.14
Conscientiousness	0.02 (0.06)	-0.14-0.13	1999-2005	.28 **	.58	.09
Neuroticism	0.00 (0.05)	-0.09-0.12	1999-2005	.59 ***	.41	.17
Openness to experience	-0.02 (0.06)	-0.17-0.11	1999-2005	.30 **	.54	.14
Greater than 30 Personality						

Table 34, cont.

Extraversion	0.01	-0.09-	1999-	.10	.80	.10
	(0.03)	0.07	2005			
Agreeableness	-0.03	-0.14-	1999-	.11	.76	-.06
	(0.04)	0.06	2005			
Conscientiousness	-0.05	-0.14-	1999-	.21 *	.80	.15
	(0.04)	0.02	2005			
Neuroticism	0.02	-0.12-	1999-	.22 *	.55	.10
	(0.06)	0.17	2005			
Openness to experience	-0.03	-0.18-	1999-	.51 ***	.34	.15
	(0.08)	0.13	2005			
Gender Difference Personality						
Extraversion	-0.12	-0.17-	1999-	-.02	.83	-.08
	(0.02)	-0.09	2005			
Agreeableness	-0.11	-0.14-	1999-	.07	.71	-.08
	(0.02)	-0.07	2005			
Conscientiousness	-0.08	-0.14-	1999-	.12	.65	-.01
	(0.02)	-0.03	2005			
Neuroticism	-0.31	-0.35-	1999-	.15	.72	.07
	(0.02)	-0.23	2005			
Openness to experience	0.08	0.01-	1999-	.34 ***	.62	-.06
	(0.02)	0.14	2005			
Age Difference Personality						
Extraversion	-0.02	-0.10-	1999-	.18 *	.48	-.05
	(0.04)	0.05	2005			
Agreeableness	0.04	-0.09-	1999-	.15	.77	.10
	(0.05)	0.17	2005			
Conscientiousness	0.07	-0.11-	1999-	.05	.51	-.05
	(0.06)	0.21	2005			
Neuroticism	-0.02	-0.11-	1999-	.15	.68	-.08
	(0.04)	0.09	2005			
Openness to experience	0.01	-0.11-	1999-	.57 ***	.37	.11
	(0.06)	0.14	2005			
Fertility Schedule						
Total fertility rate	1.95	1.63-	2010	.51 ***	.43	.02
	(0.17)	2.45				
Initiation	10.51	6.06-	2010	.19 *	.44	.13
	(1.68)	13.34				
Peak	26.83	22.72-	2010	.60 ***	.20	.23 **
	(2.49)	32.16				
Stopping	3.87	1.28-	2010	.28 **	.48	.10
	(0.82)	5.16				
Markers of Fertility Norms						
Age at first birth	24.84	22.60-	2006	.64 ***	.14	.21 *
	(1.19)	27.70				
Age at first marriage	27.31	24.45-	2010	.48 ***	.37	.08
	(1.13)	29.75				
Percent never married	30.70	24.80-	2010	.34 ***	.38	.17 *
	(2.82)	37.53				
Percent divorce	1.46	1.00-	2010	.45 ***	.60	.31 ***
	(0.22)	2.92				

Table 34, cont.

Percent cohabit	6.88 (1.09)	4.6-9.3	2010	.50 ***	.48	.07
Non-marital fertility	35.55 (6.22)	15.80-51.20	2010	.38 ***	.29	.01
Percent unintended pregnancy	51.56 (5.44)	38-65	2006	.48 ***	.38	-.01
Abortion rate	15.62 (8.35)	0.90-40.0	2008	.36 ***	.36	-.07
Family planning expenditures	107.46 (43.31)	31-245	2010	.12	.89	.05
Sociodemographic Characteristics and Value Controls						
Median income (in $1,000)	49.76 (8.06)	36.85-68.85	2010	.28 **	-	-
Percent African American	10.34 (9.55)	0.4-37.0	2010	.64 ***	-	-
Percent Hispanic	10.61 (9.88)	1.2-46.3	2010	.42 ***	-	-
Percent female	50.66 (0.75)	47.9-51.7	2010	.53 ***	-	-
Percent with a B.A.	27.16 (4.76)	17.3-38.3	2010	.35 ***	-	-
Percent urban	73.58 (14.56)	38.66-94.95	2010	.30 ***	-	-
Vote for Obama	50.51 (9.49)	32.54-71.85	2008	.18 *	-	-
Percent very religious	39.62 (8.62)	23.80-56.60	2010	.11	-	-

Note. SD = Standard Deviation. B.A. = Bachelor's degree. FEC = Federal Election Commission. Personality variables derived from data described in Rentfrow et al. (*21*). Fertility variables were derived from data described in Martin et al. (*33*). Sociodemographic characteristics were derived from the 2010 Census. Vote for Obama was derived from the Federal Election Commission (*41*) report. Percent very religious was derived from a Gallup (*42*) poll. Moran's I is a measure of spatial autocorrelation and was calculated based on a queen contiguity matrix. Residual variance refers to the proportion of variance in the main study variables remaining after sociodemographic characteristics and value variables were controlled. Residual Moran's I refers to the spatial autocorrelation of the residuals.
* $p < .05$; ** $p < .01$; *** $p < .001$.

Analytic Approach

Following the above procedure, we calculated the correlation between the aggregate personality variables and the fertility outcomes. This provides a general impression of whether individual differences in personality are associated with fertility. Because the state-level estimates are based on aggregates of thousands of individuals, the

mean estimates are very precise and typically produce robust associations (Rosnow, Rosenthal, & Rubin, 2000).

To probe whether personality factors differentially matter for fertility based on gender, we used personality aggregates derived from males and females separately. We used multiple regression to predict each fertility outcome by the estimates of male and female personality. This procedure provides an index of whether male or female personality matters more or in a different direction than personality for both sexes combined. We performed a similar approach with the two age ranges of personality, again, including both variables in a single regression. These estimates of personality tended to be correlated across gender (average r = .58) and age categories (average r = .62). This potentially introduces the problem of multicollinearity which tends to inflate standard errors and can sometimes obfuscate interpretation of the regression parameters (Cohen, Cohen, West, & Aiken, 2003). To complement the standard regression analysis, we also performed a commonality analysis (Nimon, Lewis, Kane, & Hayes, 2008). Commonality analysis partitions variance accounted for (R^2) among predictor variables into that which is unique to that predictor and that which is shared in common with the other predictors. This is accomplished by comparing the amount of variance accounted for in the outcome variable by all possible regression subsets. For our analysis based on subgroups, this entailed a comparison of three separate models predicting fertility. For example, the commonality analysis for gendered personality entailed estimating variance explained by male personality, by female personality, and by the multiple regression of male and female personality. This approach allows the overlapping variance to be identified and partitioned. Rather than treating multicollinearity as a problem to be fixed, this approach takes multicollinearity into account and provides reasonable estimates of an independent variable's effect at multiple levels.

Although our gender and age estimates of personality were moderately strongly correlated with each other, they were very strongly correlated with the estimates of personality based on the full sample. The average correlation between male and female estimates of personality and the full sample estimate was .80. For estimates based on age categories, the average correlation was .73. Therefore, we interpret common effects on fertility shared across the gender or age variables to be primarily indicative of the general association found with the full sample estimates of personality. The unique predictive power of the gender or age category variables, then, represents potential personality associations with fertility that are obscured when the full data estimates of personality are used.

Results of these analyses are presented in Table 35 under the columns labeled "Subgroup Personality." The first three columns report results for an analysis based on gender. Standardized regression parameters are reported in the first two columns. There were few associations between male or female personality and fertility. The third column reports the percent of the total variance accounted for in fertility (R^2) explained by common variance of male and female personality. Typically, the majority of variance was explained by common effects. Negative percent common variance estimates indicate suppression effects, whereby correlated individual predictors act in opposite directions, but together provide additional information. The last three columns under the "Subgroup Personality" heading report results for an analysis based on age groups. Several effects are significant, indicating that regional fertility outcomes are sensitive to levels of personality among the younger (< 30 years old) and older (≥ 30 years old) population. For example, states with higher openness among the younger population relative to the older population tended to have earlier age at first birth and marriage, and lower percent never married, cohabitation, non-marital fertility, and family planning expenditures. This

pattern of association is qualitatively different from that observed at the omnibus level. Traditional fertility practices are more prevalent in regions with greater concentrations of young people high in openness compared to older people.

To test for whether the influence of subgroup personality is relative to the personality of another subgroup, we calculated difference scores. For gendered personality, we calculated the difference between male and female personality with higher scores indicating that males tend to score higher on average in the region. For aged personality, we calculated the difference between the younger (< 30 years) and somewhat older (≥ 30 years) personality levels with higher scores indicating that the younger population tends to score higher on average in the region. We used these difference scores to correlate with the fertility outcomes.

Results of these analyses are presented in Table 35 under the columns labeled "Relative Personality." The first column reports results for an analysis based on gender, and there were relatively few significant correlations. The second column reports results for an analysis based on age, and there were substantially more significant correlations. This was particularly the case for extraversion and openness, and in reference to fertility outcomes related to family formation.

Finally, the omnibus, aggregate regional personality estimates were moderately intercorrelated (average absolute value r = .45). As a sensitivity analysis, we evaluated whether associations between personality traits and fertility were due to unique or common effects using commonality analysis. To accomplish this goal, we evaluated all possible regression subsets for the five predictor variables (i.e., univariate associations with fertility, every pairwise combination of personality traits, and all additional levels of combination including a multiple regression with all five traits simultaneously predicting the outcome).

Results of these analyses are presented in Table 35 under the columns labeled "Trait Covariation." The first column reports standardized regression parameters, and the second column reports the percent of the total R^2 explained by common variance. Common variance explained large portions of the effect for total fertility rate and non-marital fertility rate, but other outcomes frequently displayed unique associations with personality.

Table 35. Personality-fertility associations by trait, subgroup, relative trait, and model

	Subgroup Personality (multiple regression)						Relative Personality (correlation)		Trait Covariation (multiple regression)	
	β's		% R^2	β's		% R^2	Male -	< 30 -		% R^2
	Male	Female	Common	< 30	≥ 30	Common	Female	≥ 30	β's	Common
Panel A: Extraversion										
Total fertility rate	.24	.01	61.56%	.06	.24	55.96%	.13	-.03	**-.42**	42.47%
Initiation	-.16	-.07	73.59%	-.31	.18	-29.15%	-.09	-.22	-.40	5.44%
Peak	-.03	.24	51.10%	.37	-.26	-46.81%	-.03	.25	.11	17.15%
Stopping	-.15	**.56**	33.62%	**.70**	**-.42**	-35.54%	-.13	**.51**	.19	24.73%
Age at first birth	.25	-.19	-59.08%	.08	-.04	-22.11%	.16	-.13	.12	3.84%
Age at first marriage	.41	**-.59**	-45.50%	-.13	-.13	65.02%	.06	-.15	.20	6.32%
Percent never married	.08	-.25	28.30%	-.27	.10	0.28%	.05	-.12	.15	-24.56%
Percent divorce	.09	-.08	-76.00%	.19	-.19	-65.03%	-.13	.20	.09	-7.84%
Percent cohabit	.24	-.46	-9.57%	-.21	-.02	48.39%	-.08	-.21	.06	20.42%
Non-marital fertility	-.12	-.23	74.99%	**-.45**	.12	14.24%	**-.34**	**-.28**	-.13	63.31%
Percent unintended pregnancy	-.42	.11	34.35%	**-.73**	**.55**	-54.21%	-.25	**-.64**	-.45	12.12%
Abortion rate	-.27	.29	-74.70%	-.10	.16	-37.04	-.17	**-.28**	-.24	9.28%
Family planning expenditures	-.05	-.23	69.14%	**-.43**	.22	-21.83%	-.07	**-.30**	-.18	2.96%
Panel B: Agreeableness										
Total fertility rate	.34	.20	84.60%	.29	.31	63.28%	-.09	.22	.16	-
Initiation	.00	.11	74.67%	-.16	.33	-19.91%	-.05	-.13	.15	-
Peak	.02	.09	62.46%	-.01	.13	35.94%	-.04	.00	.00	-
Stopping	.07	-.03	30.05%	-.06	.19	4.38%	.02	-.05	**-.48**	-
Age at first birth	-.13	.27	22.96%	-.02	.20	33.50%	.05	.10	.22	-
Age at first marriage	-.27	.15	1.28%	.07	-.25	8.85%	-.09	.18	.23	-
Percent never married	.17	-.23	-40.98%	.12	-.26	-12.09%	-.21	.25	.46	-
Percent divorce	.09	-.08	-66.09%	-.16	.22	-51.27%	.18	-.26	-.02	-
Percent cohabit	.09	-.30	50.40%	-.28	.08	11.45%	.11	-.20	.13	-
Non-marital fertility	-.46	.19	31.39%	.05	**-.42**	28.45%	**-.36**	.17	.07	-
Percent unintended pregnancy	-.04	.10	40.86%	.20	-.19	-63.30%	-.09	.12	.37	-
Abortion rate	-.02	.10	64.29%	.05	.03	61.43%	.03	-.01	-.03	-
Family planning expenditures	.18	-.09	14.26%	.14	-.11	-57.20%	-.15	.10	**.55**	-
Panel C: Conscientiousness										

Table 35, cont.

Total fertility rate	.30	.18	60.06%	**.41**	.05	33.56%	.04	.25	.28	-
Initiation	.11	.03	52.57%	-.09	.08	-49.39%	.03	-.08	.01	-
Peak	.34	.20	-32.49%	**.37**	-.23	-36.20%	.21	**.29**	.38	-
Stopping	.22	-.03	25.22%	**.55**	**-.29**	-27.75%	.11	**.46**	**.52**	-
Age at first birth	.07	.21	54.86%	.12	.12	50.04%	-.11	-.15	.27	-
Age at first marriage	-.10	-.06	60.32%	-.07	-.11	48.06%	-.08	-.16	.01	-
Percent never married	.09	-.28	1.92%	.03	**-.36**	18.58%	.00	**.29**	-.24	-
Percent divorce	.08	-.02	12.35%	-.12	.17	-42.80%	.09	-.07	-.10	-
Percent cohabit	-.10	-.13	61.48%	-.19	.00	23.91%	-.11	**-.29**	-.31	-
Non-marital fertility	-.29	-.12	57.01%	**-.54**	.16	-3.28%	-.16	-.26	-.45	-
Percent unintended pregnancy	-.09	-.02	50.96%	-.22	.01	-12.59%	-.12	**-.41**	-.21	-
Abortion rate	-.04	.20	19.12%	.11	.05	45.06%	-.13	-.26	.20	-
Family planning expenditures	.03	-.09	6.63%	-.01	-.11	49.77%	-.18	-.08	-.25	-
Panel D: Neuroticism										
Total fertility rate	**-.38**	.10	-1.26%	**-.52**	-.10	53.38%	.12	-.04	-.20	-
Initiation	.16	.02	-0.57%	.00	-.28	40.16%	.13	.17	-.06	-
Peak	.00	.04	0.08%	-.14	**.25**	-26.68%	-.13	-.17	.02	-
Stopping	-.19	.01	0.40%	**-.39**	**.51**	-53.80%	**-.43**	**-.40**	-.07	-
Age at first birth	-.17	.05	8.23%	.08	-.18	-8.50%	.21	.19	**.42**	-
Age at first marriage	.26	-.02	33.25%	.15	.07	59.98%	.11	.13	**.39**	-
Percent never married	.23	-.04	25.16%	.00	.18	42.40%	.03	.18	.09	-
Percent divorce	-.13	.05	-3.51%	.09	-.17	-26.29%	-.14	-.14	-.05	-
Percent cohabit	.03	.05	59.89%	.17	-.07	-11.46%	.05	.10	**.32**	-
Non-marital fertility	.39	-.16	-7.37%	.29	-.11	-1.19%	.08	**.33**	-.01	-
Percent unintended pregnancy	**.43**	**-.52**	-59.20%	.18	-.36	-22.58%	**.45**	**.48**	.15	-
Abortion rate	-.08	-.02	56.03%	.07	-.17	-6.09%	.21	.21	**.35**	-
Family planning expenditures	.28	-.29	-63.36%	.17	-.24	-49.46%	**.31**	**.33**	.02	-
Panel E: Openness										
Total fertility rate	-.01	**-.53**	50.54%	-.13	**-.49**	62.41%	**.31**	.17	**-.50**	-
Initiation	.25	-.23	-68.42%	.04	-.10	-7.96%	.17	.06	-.20	-
Peak	-.31	.26	-64.59%	.08	-.06	-65.26%	-.19	-.05	.35	-
Stopping	-.14	-.14	70.12%	-.16	-.09	68.06%	.06	-.02	-.08	-
Age at first birth	-.10	.06	-29.35%	-.14	.13	-68.07%	-.24	**-.28**	.18	-
Age at first marriage	-.15	.32	-12.45%	-.06	.31	30.56%	**-.35**	**-.37**	.19	-
Percent never married	.00	.34	49.65%	.00	**.42**	49.16%	-.09	**-.30**	**.60**	-

Table 35, cont.

Percent divorce	-.04	.03	-45.96%	.13	-.19	-50.62%	.12	.17	-.03	-
Percent cohabit	-.11	.35	13.90%	.26	-.01	45.81%	**-.34**	**-.32**	.14	-
Non-marital fertility	-.11	.34	12.36%	.12	.14	70.11%	-.25	**-.31**	-.11	-
Percent unintended pregnancy	.03	.00	46.22%	-.12	.16	-56.82%	-.13	-.08	-.18	-
Abortion rate	-.12	-.08	69.36%	-.23	.05	27.82%	-.14	-.19	-.26	-
Family planning expenditures	.16	.10	68.70%	-.05	.30	35.32%	-.11	**-.31**	.31	-

Note. The second and third columns report standardized regression coefficients (β) from a multiple regression using male and female personality. The fourth column reports the percent of the total effect due to common variance of male and female personality. The fifth through sixth column report similar multiple regression results for personality age groups. The seventh and eighth column report correlations between relative personality (i.e., male - female personality and young - old personality) and fertility. The ninth column reports multiple regression results from models that included each personality trait to control for trait covariation. A total of 13 models were fit, and parameters are listed separately for ease of presentation. The final column reports the percent of the total effect explained by variance common to all personality traits. Parameters printed in bold are significant at $p < .05$. Negative % Common values indicate a suppressor effects. All variables adjusted for sociodemographic characteristics and value controls.

References

Achenbach, T.M. (1991). *Manual for the Child Behavior Checklist/4-18*. Burlington, VT: University of Vermont, Department of Psychiatry.

Alexander, K. L., Entwisle, D. R., & Bedinger, S. D. (1994). When expectations work: Race and socioeconomic differences in school performance. *Social Psychology Quarterly, 57,* 283-299. doi: 10.2307/2787156

Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt, Rinehart, & Winston.

Alvergne, A., Jokela, M., & Lummaa, V. (2010). Personality and reproductive success in a high-fertility human population. *Proceedings of the National Academy of Sciences of the United States of America, 107,* 11745-11750. doi: 10.1073/pnas.1001752107

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: Elsevier.

Andrew, M. & Hauser, R. M. (2011). Adoption? Adaptation? Evaluating the formation of educational expectations. *Social Forces, 90,* 497-520. doi: 10.1093/sf/sor005

Anselin, L. (1988). *Spatial econometrics: Methods and models*. Kluwer Academic Publishers: Dordrecht, The Netherlands.

Anselin, L., Bera, A. K., Floras, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics, 26,* 77-104.

Aplin, L. M., Farine, D. R., Morand-Ferron, J., Cole, E. F., Cockburn, A., & Sheldon, B. C. (2013). Individual personalities predict social behavior in wild networks of great tits. *Ecology Letters.* doi: 10.1111/ele.12181

Armstrong, P. I. & Anthoney, S. F. (2009). Personality facets and RIASEC interests: An integrated model. *Journal of Vocational Behavior, 75,* 346-359.

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

Ashton, M. C. (1998). Personality and job performance: The importance of narrow traits. *Journal of Organizational Behavior, 19,* 289-303.

Ashton, M. C., Lee, K., Goldberg, L. R., de Vries, R. E. (2009). Higher-order factors of personality: Do they exist? *Personality and Social Psychology Review, 13,* 79-91.

Avinun, R., & Knafo, A. (2013, August). Parenting as a reaction evoked by children's genotype: A meta-analysis of children-as-twins studies. *Personality and Social Psychology Review*. Advanced online publication. doi: 10.1177/1088868313498308

Avinun, R., Ebstein, R. P., & Knafo, A. (2012). Human maternal behavior is associated with arginine vasopressin receptor 1A gene. *Biology Letters, 8,* 894-896. doi: 10.1098/rsbl.2012.0492

Bäckström, M., Björklund, F., & Larsson, M. R. (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality, 43,* 335-344.

Bakermans-Kranenburg, M. J., & van IJzenboorn, 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

Bakermans-Kranenburg, M., & Van IJzendoorn, M. (2008). Oxytocin receptor (OXTR) and serotonin transporter (5-HTT) genes associated with observed parenting. *Social Cognitive and Affective Neuroscience, 3,* 128.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

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: Guilford Press.

Barnett, W. S. (1995). Long-term effects of early childcare programs on cognitive and school outcomes. *The Future of Children, 5,* 25-50. doi: 10.2307/1602366

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173-1182. doi: 10.1037/0022-3514.51.6.1173

Barrick, M. R. & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*, 1-26.

Bayley, N. (1969). *Manual for the Bayley Scales of Infant Development*. New York: Psychological Corporation.

Beal, S. J. & Crockett, L. J. (2010). Adolescents' occupational and educational aspirations and expectations: Links to high school activities and adult educational attainment. *Developmental Psychology, 46,* 258-265. doi: 10.1037/a0017416

Becker, G. S. (1981). *A treatise on the family*. Cambridge, MA: Harvard University Press.

Bell, R. Q. (1968). A reinterpretation of the direction of effects in studies of socialization. *Psychological Review, 75,* 81-95. doi: 10.1037/h0025583

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

Bentler, P. M. & Satorra, A. (2010). Testing model nesting and equivalence. *Psychological Methods*, *15*, 111-123.

Berg, V., Rotkirch, A., Väisänen, H., & Jokela, M. (2013). Personality is differentially associated with planned and non-planned pregnancies. *Journal of Research in Personality, 47,* 296-305. doi: 10.1016/j.jrp.2013.01.010

Bergeman, C. S., Chipuer, H. M., Plomin, R., Pedersen, N. L., McClearn, G. E., Nesselroade, J. R., . . . McCrae, R. R. (1993). Genetic and environmental effects

on openness to experience, agreeableness, and conscientiousness: An adoption/twin study. *Journal of Personality*, *61*, 159-179.

Billari, F. C., & Kohler, H.-P. (2004). Patterns of low and lowest-low fertility in Europe. *Population Studies: A Journal of Demography, 58,* 161-176.

Bishop, B. (2009). *The big sort: Why the clustering of like-minded America is tearing us apart.* New York: Houghton Mifflin.

Bivand, R. (2014). *Spdep: Spatial dependence: Weighting schemes, statistics and models. R package version 0.5-77.* Available at Available at http://cran.r-project.org/web/packages/spdep/spdep.pdf

Blau, P. M., & Duncan, O. D. (1967). *The American occupational structure*. New York: John Wiley and Sons.

Bleeker, M. M. & Jacobs, J. E. (2004). Achievement in math and science: Do mothers' beliefs matter 12 years later? *Journal of Educational Psychology, 96,* 97-109. doi: 10.1037/0022-0663.96.1.97

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

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, *117*, 187-215.

Block, J. (2010). The Five-Factor framing of personality and beyond: Some ruminations. *Psychological Inquiry, 21,* 2-25.

*Blonigen, D. M., Carlson, M. D., Hicks, B. M., Kreuger, 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

Bloom, D. E., Canning, D., Fink, G., & Finlay, J. E. (2010). The cost of low fertility in Europe. *European Journal of Population, 26,* 141-158.

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

Bongaarts, J., & Watkins, S. C. (1996). Social interactions and contemporary fertility transitions. *Population and Development Review, 22,* 639-682.

Bornstein, M. H., Hahn, C.-S., & Suwalsky, J. T. D. (2013, August). Physically developed and exploratory young infants contribute to their own long-term

academic achievement. *Psychological Science*. Advanced online publication. doi: 10.1177/0956797613479974

Bornstein, M. H., Hahn, C.-S., & Wolke, D. (2013). Systems and cascades in cognitive development and academic achievement. *Child Development, 84,* 154-132. doi: 10.1111/j.1467-8624.2012.01849.x

Bouchard, T. J., Jr. (1993). Experience producing drive theory: How genes drive experience and shape personality. *Acta Paediatrica Supplement, 422,* 60-64. doi: 10.1111/j.1651-2227.1997.tb18347.x

Bouchard, T. J., Jr. (1997). The genetics of personality. In K. Blum and E. P. Noble (Eds.), *Handbook of Psychiatric Genetics* (pp. 273-296). Boca Raton, FL: Circle Press, Inc.

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: Basic Books.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53,* 371-399. Doi: 10.1146/annurev.psych.53.100901.135233

*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

Briley, D. A., & Tucker-Drob, E. M. (2013). Explaining the increasing heritability of cognition over development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science, 24,* 1704-1713. doi: 10.1177/0956797613478618

Briley, D. A., & Tucker-Drob, E. M. (2014). Genetic and environmental continuity in personality development: A meta-analysis. *Psychological Bulletin, 140,* 1303-1331.

Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review, 101*, 568–586. DOI: 10.1037/0033-295X.101.4.568

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.

Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review, 106,* 676-713.

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

Carver, C. S., Sutton, S. K., & Scheier, M. F. (2000). Action, emotion, and personality: Emerging conceptual integration. *Personality and Social Psychology Bulletin*, *26*, 741-751.

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., 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., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., … & Moffitt, T. E. (2014). The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science, 2,* 119-137. doi: 10.1177/2167702613497473

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology, 56,* 453-484.

Cherlin, A. J. (2010). Demographic trends in the Unites States: A review of research in the 2000s. *Journal of Marriage and Family, 72,* 403-419. doi: 10.1111/j.1741-3737.2010.00710.x

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

Choo, E., & Siow, A. (2006). Who marries whom and why. *Journal of Political Economy, 114,* 175-201.

Christiansen, N. D. & Robie, C. (2011). Further consideration of the use of narrow trait scales. *Canadian Journal of Behavioral Sciences, 43,* 183-194.

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

Cleland, J., & Wilson, C. (1987). Demand theories of the fertility transition: An iconoclastic view. *Population Studies, 41,* 5-30.

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.

Coale, A. J., & Watkins, S. C. (1986). *The decline of fertility in Europe: The revised proceedings of a conference on the Princeton European Fertility Project.* Princeton: Princeton University Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Erlbaum.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Collins, W. A., Maccoby, E. E., Steinberg, L., Hetherington, E. M., & Bornstein, M. H. (2010). Contemporary research on parenting: The case for nature and nurture. *American Psychologist, 55,* 218-232. DOI: 10.1037/0003-066X.55.2.218

Colman, S., & Joyce, T. (2011). Regulating abortion: Impact on patients and providers in Texas. *Journal of Policy Analysis and Management, 30,* 775-797. doi: 10.1002/pam.20603

Conger, R. D., & Donnellan, M. B. (2007). An interactionist perspective on the socioeconomic context of human development. *Annual Review of Psychology, 58,* 175-199. doi: 10.1146/annurev.psych.58.110405.085551

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

Corff, Y. L. & Toupin, J. (2010). The five-factor model of personality at the facet level: Association with antisocial personality disorder symptoms and prediction of antisocial behavior. *Journal of Psychopathological Behavior Assessment, 32,* 586-594.

Costa, P. T., & 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. (1992a). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr. & McCrae, R. R. (1992b). Four ways the five factors are basic. *Personality and Individual Differences, 13,* 653-665.

Costa, P. T., Jr. & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment, 64,* 21-50.

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

Costa, P. T., Jr., McCrae, R. R., & Holland, J. L. (1984). Personality and vocational interests in an adult sample. *Journal of Applied Psychology*, *69*, 390-400.

Costa, P. T., Jr., McCrae, R. R., Dembroski, T. M. (1989). Agreeableness vs. antagonism: Explication of a potential risk factor for CHD. In A. Siegman & T. M. Dembroski (Eds.), *In search of coronary-prone behavior: Beyond Type A* (pp. 41-63). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cote, J., Dreiss, A., & Clobert, J. (2008). Social personality trait and fitness. *Proceedings of the Royal Society B-Biological Sciences, 275,* 2851-2858.

Courtiol, A., Pettay, J. E., Jokela, M., Rotkirch, A., & Lummaa, V. (2012). Natural and sexual selection in a monogamous historical human population. *Proceedings of the National Academy of Science, 109,* 8044-8049. doi: 10.1073/pnas.1118174109

Cronbach, L . J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302.

Cronbach, L. J. & Gleser, G. C. (1957). *Psychological tests and personnel decisions*. Urbana, IL: University of Illinois Press.

Cronbach, L. J. (1960). *Essentials of Psychological Testing* (2nd ed.). Oxford, England: Harper.

Crosnoe, R. & Johnson, M. K. (2011). Research on adolescence in the twenty-first century. *Annual Review of Sociology, 37,* 439-460. doi: 10.1146/annurev-soc-081309-150008

Cuijpers, P., Smit, F., Penninx, B. W. J. H., de Graaf, R., ten Have, M., & Beekman, A. T. F. (2010). Economic costs of neuroticism: A population-based study. *Archives of General Psychiatry, 67,* 1086-1093.

D'Addio, A. C., & d'Ercole, M. (2005). *Policies, institutions and fertility rates: A panel data analysis for OECD countries.* Organization for Economic Cooperation and Development, Paris, France.

D'Onofrio, B. M., Lahey, B. B., Turkheimer, E., & Lichtenstein, P. (2013). Critical need for family-based, quasi-experimental designs in integrating genetic and social science research. *American Journal of Public Health, 103,* S46-S55. doi: 10.2105/AJPH.2013.301252

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). The return of Phineas Gage: Clues about the brain from the skull of a famous patient. *Science, 264,* 1102-1105.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology, 19,* 294-304. doi: 10.1037/0893-3200.19.2.294

Day, J. C. & Newburger, E. C. (2002). *The big payoff: Educational attainment and synthetic estimates of work-life earnings* (Current Population Report P23-210). Washington, D.C.: U.S. Census Bureau.

*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. (2010). Meta-analysis of genome-wide association studies for personality. *Molecular Psychiatry, 17,* 337-349. doi:10.1038/mp.2010.128

de Moor, M. M., Costa, P. T., Jr., Terracciano, A., Krueger, R. F., de Gues, E. J., Toshiko, T., . . . Boomsma, D. I. (*in press*). Meta-analysis of genome-wide association studies for personality. *Molecular Psychiatry*.

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest, 11,* 53-79. doi: 10.1177/1529100610387081

DeLamater, J. (1981). The social control of sexuality. *Annual Review of Sociology, 7,* 263-290.

Demeny, P. (2003). Population policy dilemmas in Europe at the dawn of the twenty-first century. *Population and Development Review, 29,* 1-28.

DeYoung, C. G. (2006). Higher-order factors of the big five in a multi-informant sample. *Journal of Personality and Social Psychology, 91,* 1138-1151.

DeYoung, C. G., Hasher, L., Djikic, M. Criger, B., & Peterson, J. B. (2007). Morning people are stable people: Circadian rhythm and the higher-order factors of the big five. *Personality and Individual Differences, 43,* 267-276.

DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris, X., Rajeevan, N., & Gray, J. R. (2010). Testing predictions from personality neuroscience: Brain structure and the big five. *Psychological Science, 21,* 820-828. doi: 10.1177/0956797610370159

DeYoung, C. G., Peterson, J. B., Séguin, J. R., & Tremblay, R. E. (2008). Externalizing behavior and the higher order factors of the big five. *Journal of Abnormal Psychology, 117,* 947-953.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93,* 880-896.

Dick, D. M. (2011). Gene-environment interaction in psychological traits and disorders. *Annual Review of 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 and Drugs, 2,* 185-194.

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108,* 346-369.

Digman, J. M. & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis, comparison, and interpretation of six major studies. *Multivariate Behavioral Research*, *16*, 149-170.

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

Digman, J. M. (1997). Higher order factors of the Big Five. *Journal of Personality and Social Psychology, 73,* 1246-1256.

DiPrete, T. A., & Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research, 41,* 1-15. doi: 10.1016/j.ssresearch.2011.09.001

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,* 237-263. DOI: 10.1111/j.1467-6494.2007.00438.x

Downey, D. B., von Hippel, P. T., Broh, B. A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school years. *American Sociological Review, 69,* 613-635.

Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of Conscientiousness in the prediction of job performance: Examining the intercorrelations and incremental validity of narrow traits. *Journal of Applied Psychology, 91,* 40-57.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., …& Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43,* 1428-46.

Duncan, O. D., Featherman, D. L., & Duncan, B. (1972). *Socioeconomic background and achievement.* New York: Seminar Press.

Easterlin, R. A., & Crimmins, E. M. (1985). *The fertility revolution: A supply-demand analysis.* Chicago, IL: University of Chicago Press.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53,* 109-132. doi: 10.1146/annurev.psych.53.100901.135153

Ekehammar, B. & Akrami, N. (2007). Personality and prejudice: From big five personality factors to facets. *Journal of Personality, 75,* 899-926.

Erdle, S. & Rushton, J. P. (2010). The general factor of personality, BIS – BAS, expectancies of reward and punishment, self-esteem, and positive and negative affect. *Personality and Individual Differences, 48,* 762-766.

Erdle, S., Gosling, S. D., & Potter, J. (2009). Does Self-Esteem account for the higher-order factors of the Big Five? *Journal of Research in Personality, 43,* 921-922.

Erdle, S., Irwing, P., Rushton, J. P., & Park, J. (2010). The general factor of personality and its relation to self-esteem in 628, 640 internet respondents. *Personality and Individual Differences, 48,* 343-346.

*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: Norton. (Original work published 1950)

Eysenck, H. J. (1992a). A reply to Costa and McCrae. P or A and C- the role of theory. *Personality and Individual Differences, 13,* 867-868.

Eysenck, H. J. (1992b). Four ways five factors are *not* basic. *Personality and Individual Differences*, *13*, 667-673. doi: 10.1016/1091-8869(92)90237-J

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach.* New York: Plenum Press.

Eysenck, H. J., & Eysenck, S. B. (1975). *Manual of the Eysenck Personality Questionnaire*. San Diego, CA: Educational and Industrial Testing Service.

Eysenck, H. J. & Eysenck, S. B. G. (1976). *Psychoticism as a dimension of personality*. New York: Crane, Russak, & Co.

Faraone, S. V., Tsuang, M. T., & Tsuang, D. W. (1999). *Genetics of mental disorders: A guide for students, clinicians, and researchers*. New York: Guilford Press.

Federal Election Commission (2009, July). *Federal elections 2008*. Retrieved from http://www.fec.gov/pubrec/fe2008/federalelections2008.shtml

Fein, E. C. & Klein, H. J. (2011). Personality predictors of behavioral self-regulation: Linking behavioral self-regulation to five-factor model factors, facets, and a compound trait. *International Journal of Selection and Assessment, 19,* 132-144.

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

Ferguson, E., Chamorro-Premuzic, T., Pickering, A., & Weiss, A. (2011). Five into one doesn't go: A critique of the general factor of personality. In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbook of individual differences* (pp.162-186). West Sussex, UK: Wiley-Blackwell.

Finer, L. B., & Kost, K. (2011). Unintended pregnancy rates at the state level. *Perspectives on Sexual and Reproductive Health, 43,* 78-87. doi: 10.1363/4307811

Forget-Dubois, N., Pérusse, D., Turecki, G., Girard, A., Billette, J., Rouleau, G., …Tremblay, R. E. (2003). Diagnosing zygosity in infant twins: Physical similarity, genotyping, and chorionicity. *Twin Research, 6,* 479-485. doi: 10.1375/136905203322686464

*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

Fredricks, J. A. & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental Psychology, 38,* 519-533. doi: 10.1037/0012-1649.38.4.519

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: Hogarth Press. (Original work published 1908)

Frome, P. M. & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology, 74,* 435-452. doi: 10.1037/0022-3514.74.2.435

Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception*. San Diego, CA: Academic Press.

Funder, D. C. (2001). Personality. *Annual Review of Psychology, 52*, 197-221.

*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

Gallup (2010). *State of the states*. Retrieved from http://www.gallup.com/poll/125066/state-states.aspx

*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

Glass, J., & Levchak, P. (2014). Red states, blue states, and divorce: Understanding the impact of conservative Protestantism on regional variation in divorce rates. *American Journal of Sociology,119,* 1002-1046.

Gold, R. B. (2002). Hierarchy crackdown clouds future of sterilization, EC provision at Catholic hospitals. *Guttmacher Report on Public Policy, 5*.

Goldberg, L. R. (1993). The structure of personality traits: Vertical and horizontal aspects. In D. C. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds.), *Studying lives through time: Personality and development* (pp. 169-188). Washington, D. C.: American Psychological Association.

Goldsmith, H. H. (1996). Studying temperament via construction of the Toddler Behavior Assessment Questionnaire. *Child Development, 67,* 218–235. DOI: 10.1111/j.1467-8624.1996.tb01730.x

Gonzalez-Pienda, J. A., Nunez, J. C., Gonzalez-Pumariega, S., Alvarez, L., Roces, C., & Garcia, M. (2002). A structural equation model of parental involvement, motivational and aptitudinal characteristics, and academic achievement. *The Journal of Experimental Education, 70,* 257-287. doi:10.1080/00220970209599509

Goodwin, P., McGill, B., & Chandra, A. (2009). Who marries and when? Age at first marriage in the United States: 2002. *NCHS Data Brief,* No. 19.

Gordon, L. (2002). *The moral property of women: A history of birth control politics in America*. Champaign-Urbana, IL: University of Illinois Press.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist, 59,* 93-104.

Gough, H. B. (1957). *Manual for the California Psychological Inventory*. Palo Alto: Consulting Psychologists Press.

Gough, H. B., & Bradley, P. (1996). *CPI manual* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press.

Gough, H. B., & Heilbrun, A. B. (1983). *The Adjective Check List manual* (2nd ed.). Palo Alto, CA: Consulting Psychologists Press.

Grammich, C., DaVanzo, J., & Stewart, K. (2004). Changes in American opinion about family planning. *Studies in Family Planning, 35,* 197-206.

Grammich, C., Hadaway, K., Houseal, R., Jones, D. E., Krindatch, A., Stanley, R., & Taylor, R. H. (2012). *U.S. Religion Census: Religious congregations & membership study*. Association of Statisticians of American Religious Bodies.

Gurven, M., von Rueden, C., Stieglitz, J., Kaplan, H., Rodriguez, D. E. (2013, October). The evolutionary fitness of personality traits in a small-scale subsistence society. *Evolution and Human Behavior*. Advanced online publication. doi: 10.1016/j.evolhumbehav.2013.09.002

Hanson, S. L. (1994). Lost talent: Unrealized educational aspirations and expectations among U.S. youths. *Sociology of Education, 67,* 159-183. doi: 10.2307/2112789

Harackiewicz, J. M., Rozek, C. S., Hulleman, C. S., & Hyde, J. S. (2012). Helping parents to motivate adolescents in mathematics and science: An experimental test

of a utility-value intervention. *Psychological Science, 23,* 899-906. doi: 10.1177/0956797611435530

Harden, K. P. (2014). Genetic influences on adolescent sexual behavior: Why genes matter for environmentally oriented researchers. *Psychological Bulletin, 140,* 434-465. doi: 10.1037/a0033564

Harper, S. (2014). Economic and social implications of aging societies. *Science, 346,* 587-591.

Harper, C. C., Henderson, J. T., & Darney, P. D. (2005). Abortion in the United States. *Annual Review of Public Health, 26,* 501-512. doi: 10.1146/annurev.publhealth.26.021304.144351

Haworth, C. M. A., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J. C., van Beijsterveldt, C. E. M., … Plomin, R. (2010). The heritability of cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry, 15,* 1112-1120. doi: 10.1038/mp.2009.55

Hayes, K. J. (1962). Genes, drives, and intellect. *Psychological Reports, 10,* 299-342. doi: 10.2466/pr0.1962.10.2.299

Heath, A. C. & Martin, N. G. (1990). Psychoticism as a dimension of personality: A multivariate genetic test of Eysenck and Eysenck's Psychoticism construct. *Journal of Personality and Social Psychology*, *58*, 111-121.

Heath, A. C., Berg, K., Eaves, L. J., Solaas, M. H., Corey, L. A., Sundet, J., … Nance, W. E. (1985). Education policy and the heritability of educational attainment. *Nature, 314,* 734-736. doi:10.1038/314734a0

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312,* 1900-1902.

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

Hirsh, J. B., DeYoung, C. G., & Peterson, J. B. (2009). Metatraits of the big five differentially predict engagement and restraint of behavior. *Journal of Personality, 77(4),* 1-17.

Hofstee, W. K. B., de Raad, B., & Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, *63*, 146-163.

*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

Hutteman, R., Bleidorn, W., Penke, L., & Denissen, J. J. A. (2013). It takes two: A longitudinal dyadic study on predictors of fertility outcomes. *Journal of Personality, 81,* 487-498. doi: 10.1111/jopy.12006

Ingraham, C. (2014). Interactive map: The oldest and youngest congressional delegations. *The Washington Post*.

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., Jonkmann, 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.

Jacob, B. A. & Linkow, T. W. (2011). Educational expectations and attainment. In G. J. Duncan, & R. J. Murnane (Eds.), *Wither opportunity?* (pp. 133-162). New York: Russell Sage Foundation.

Jacobs, J. E. & Eccles, J. S. (2000). Parents, task values, and real-life achievement-related choices. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic motivation* (pp. 405-439). San Diego, CA: Academic Press.

Jacobs, J. E., & Eccles, J. S. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of Personality and Social Psychology, 63,* 932-944. doi: 10.1037/0022-3514.63.6.932

Jang, K. L., Livesley, W. J., & Vernon, P. A. (1996). Heritability of the Big Five personality dimensions and their facets: A twin study. *Journal of Personality*, *64*, 577-591.

Jang, K. L., Livesley, W. J., Angleitner, A., Riemann, R., & Vernon, P. A. (2002). Genetic and environmental influences on the covariance of facets defining the domains of the Five-Factor model of personality. *Personality and Individual Differences*, *33*, 83-101.

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

Jang. K. L., Livesley, W. J., Ando, J., Yamagata, S., Suzuki, A., Angleitner, … Spinath, F. (2006). Behavioral genetics of the higher-order factors of the Big Five. *Personality and Individual Differences*, *41*, 261-272.

John, O. P. & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-139). New York: Guilford Press.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory – versions 4a and 54*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.

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: 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., & Krueger, R. F. (2004). Genetic and environmental structure of adjectives describing the domains of the Big Five Model of personality: A nationwide US twin study. *Journal of Research in Personality*, *38*, 448-472.

*Johnson, W., McGue, M., & Krueger, R. F. (2005). Personality stability in late adulthood: A behavior genetic analysis. *Journal of Personality, 73,* 523-551. 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

Johnson-Hanks, J. A., Bachrach, C. A., Morgan, S. P., & Kohler, H.-P. (2011). *Understanding family change and variation: Toward a theory of conjunctural action.* New York: Springer.

Jokela, M. (2012). Birth-cohort effects in the association between personality and fertility. *Psychological Science, 23,* 835-841.

Jokela, M., Alvergne, A., Pollet, T. V., & Lummaa, V. (2011). Reproductive behavior and personality traits of the five factor model. *European Journal of Personality, 25,* 487-500.

Jokela, M., Alvergne, A., Rotkirch, A., Rickard, I. J., & Lummaa, V. (2013). Associations between family size and offspring education depend on aspects of parental personality. *Personality and Individual Differences, 58,* 95-100.

Jokela, M., Kivimäki, M., Elovainio, M., & Keltikangas-Järvinen, L. (2009). Personality and having children: A two-way relationship. *Journal of Personality and Social Psychology, 96,* 208-230. doi: 10.1037/a0014058

Jones, R. K., & Kooistra, K. (2011). Abortion incidence and access to service in the United States, 2008. *Perspectives on Sexual and Reproductive Health, 43,* 41-50. doi: 10.1363/4304111

Jost, J. T., Glaser, J., Kruglanski, A. W., Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin, 129,* 339-375.

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, …& 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/s10519-005-6141-7

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

Kim, H. & Kim, H. (2011). Trends and directions in personality genetic studies. *Genomics & Infomatics, 9,* 45-51.

Kirk, D. (1996). Demographic transition theory. *Population Studies, 50,* 361-387.

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. *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: 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: Lawrence Erlbaum Associates.

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,* 1485-1521.

*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

Lerner, R. M., & Busch-Rossnagel, N. A. (1981). *Individuals as producers of their development: A life-span perspective*. New York: Academic Press.

LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Taylor & Francis: Boca Raton, LA.

Lesthaeghe, R. (1983). A century of demographic and cultural change in Western Europe: An exploration of underlying dimensions. *Population and Development Review, 9,* 411-435.

Lesthaeghe, R. (2014). The second demographic transition: A concise overview of its development. *Proceedings of the National Academy of Sciences of the United States of America, 111,* 18112-18115.

Lesthaeghe, R., & Moors, G. (1995). Living arrangements, socio-economic positions and values among young adults: A pattern description for France, Germany, Belgium and the Netherlands, 1981-1990. In D. Coleman (ed.), *Europe's population in the 1990s*. Oxford: Oxford University Press.

Lesthaeghe, R., & Neels, K. (2002). From the first to the second demographic transition: An interpretation of demographic innovation in France, Belgium, and Switzerland. *European Journal of Population, 18,* 325-360.

Lesthaeghe, R., & Neidert, L. (2006). The second demographic transition in the United States: Exception or textbook example? *Population and Development Review, 32,* 669-698.

Lesthaeghe, R., & Neidert, L. (2009). US Presidential elections and the spatial pattern of the American second demographic transition. *Population and Development Review, 35,* 391-400.

Lesthaeghe, R., & Surkyn, J. R. (2002). Value orientations and the second demographic transition (SDT) in Northern, Western and Southern Europe: An update. *Demographic Research, 3,* 45-86.

Lesthaeghe, R., & van de Kaa, D. J. (1986). Twee demografische transities? (pp. 9-24). In R. J. Lesthaeghe and D. J. van de Kaa (eds.), *Bevolking: Groei en Krimp. Mens en Maatschappij*. Deventer: Van Loghum-Slaterus.

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*. Newbury Park, CA: Sage.

Loehlin, J. C. (1996). The Cholesky approach: A cautionary note. *Behavioral Genetics*, *26*, 65-69.

Loehlin, J. C., & Nichols, R. C. (1976). *Heredity, environment and personality*. Austin: University of Texas.

Lofquist, D., Lugaila, T., O'Connell, M., & Feliz, S. (2012). Households and Families: 2010. *2010 Census Briefs*.

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, 4,* 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_2 receptor exon 8 and GABA(A) alpha 6 receptor variants: An explorative study. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 141B,* 167-172. DOI: 10.1002/ajmg.b.30255

Luciano, M., Wainwright, M. A., Wright, M. J., & Martin, N. G. (2006). The heritability of conscientiousness facets and their relationship to IQ and academic achievement. *Personality and Individual Differences, 40,* 1189-1199.

Lugo-Gil, J., Tamis-LeMonda, C. A. (2008). Family resources and parenting quality: Links to children's cognitive development across the first 3 years. *Child Development, 79,* 1065-1085. doi: 10.1111/j.1467-8624.2008.01176.x

Luijk, M. P. C. M., Roisman, G. I., Haltigan, J. D., Tiemeier, H., Booth-LaForce, C., van IJzendorn, 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

Lykken, D. T. (1968). Statistical significance in psychology research. *Psychological Bulletin*, *70*, 151-159.

Maher, B. (2008). Personal genomes: The case of the missing heritability. *Nature, 456,* 18-21. doi: 10.1038/456018a.

Manning, J. E. (2011). Membership of the 112th congress: A profile. *Congressional Research Service*, *7-5700*.

Marsh, H. W. & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81,* 59-77. doi: 10.1348/000709910X503501

Marsh, H. W. & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. *Personality and Social Psychology Bulletin, 34,* 542-552. doi: 10.1177/0146167207312313

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22,* 471-492.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interests, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76,* 397-416. doi: 10.1111/j.1467-8624.2005.00853.x

Martin, J. A., Hamilton, B. E., Ventura, S. J., Osterman, M. J. K., Wilson, E. C., & Mathews, T. J. (2012). Births: Final data for 2010. *National Vital Statistics Reports, 61,* 1-72.

Mathews, T. J., & Hamilton, B. E. (2009). Delayed childbearing: More women are having their first child later in life. *NCHS Data Brief,* No. 21.

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 behavioural and emotional traits. *Behavior Genetics, 43,* 314-328. doi: 10.1007/s10519-013-9591-0

McArdle, J. J. & Goldsmith, H. H. (1990). Alternative common factor models for multivariate biometric analyses. *Behavior Genetics*, *20*, 569-608.

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. J., Jr. (1985). Comparison of EPI and psychoticism scales with measures of the five-factor model of personality. *Personality and Individual Differences, 6,* 587-597.

McCrae, R. R. (1994). Openness to experience as a basic dimension of personality. *Imagination, Cognition and Personality, 13,* 39-55.

McCrae, R. R., & Costa, P. J., Jr. (1991). Adding *Liebe und Arbeit*: The full five-factor model and well-being. *Personality and Social Psychology Bulletin*, *17*, 227-232.

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. J., Jr. (1997). Personality trait structure as human universal. *American Psychologist*, *52*, 509-516.

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: Lawrence Erlbaum Associates.

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: Guilford Press.

McCrae, R. R., Costa, P. J., Jr., & Piedmont, R. L. (1993). Folk concepts, natural language, and psychological constructs: The California Psychological Inventory and the five-factor model. *Journal of Personality*, *61*, 1-26.

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(2),* 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

McCrae, R. R., Jang, K. L., Livesley, W. J., Riemann, R., & Angleitner, A. (2001). Sources of structure: Genetic, environmental, and artifactual influences on the covariation of personality traits. *Journal of Personality*, *69*, 511-535.

McCrae, R. R., Yamagata, S., Jang, J. L., Riemann, R., Ando, J., Ono, Y., Angleitner, A., & Spinath, F. M. (2008). Substance and artifact in the higher-order Big Five. *Journal of Personality and Social Psychology, 95,* 442-455.

McDonald, P. (2000). Gender equality, social institutions and the future of fertility. *Journal of Population Research, 17,* 1-16.

*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

McGue, M., & Bouchard, T. J., Jr. (1984). Adjustment of twin data for the effects of age and sex. *Behavior Genetics, 14,* 325-343. doi: 10.1007/BF01080045

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist, 53,* 185-204. Doi: 10.1037/0003-066X.53.2.185

McWhirter, E. H. (1997). Perceived barriers to education and career: Ethnic and gender differences. *Journal of Vocational Behavior, 50,* 124-140. doi: 10.1006/jvbe.1995.1536

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.

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

Mickelson, R. A. (1990). The attitude-achievement paradox among black adolescents. *Sociology of Education, 63,* 44-61. doi: 10.2307/2112896

Miller, T. (1991). The psychotherapeutic utility of the five-factor model of personality: A clinician's experience. *Journal of Personality Assessment*, *57*, 415-433.

Miller, W. B. (1992). Personality traits and developmental experiences as antecedents of childbearing motivation. *Demography, 29,* 265-285.

Mills, M., Rindfuss, R. R., McDonald, P., & te Velde, E. (2011). Why do people postpone parenthood? Reasons and social policy incentives. *Human Reproduction Update, 17,* 848-860.

Mischel, W. (1968). *Personality and Assessment.* New York: Wiley.

Mitchell, J. T., Kimbrel, N. A., Hundt, N. E., Cobb, A. R., Nelson-Gray, R. O., Lootens, C. M. (2007). An analysis of reinforcement sensitivity theory and the five-factor model. *European Journal of Personality, 21,* 869-887.

Moffitt, T. E., Arseneault, L., Belskey, D., Dickson, N., Hancox, R. J., … & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences, 108,* 2693-2698. doi: 10.1073/pnas.1010076108

Montez, J. K., Hummer, R. A., Hayward, M. D., Woo, H., & Rogers, R. G. (2011). Trends in the educational gradient of U.S. adult mortality from 1986 through 2006 by race, gender, and age group. *Research on Aging, 33,* 145-171. doi: 10.1177/0164027510392388

Mood, A. M. (1969). Macro-analysis of the American educational system. *Operations Research, 8,* 770-784.

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika, 37,* 17-23.

Morgan, S. P. (1996). Characteristic features of modern American fertility. *Population and Development Review, 22,* 19-63.

Morgan, S. P. (2010). Is low fertility a twenty-first century demographic crisis? *Demography, 40,* 589-603.

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: Lawrence Erlbaum Associates.

Murray, H. A. (1938). *Explorations in personality: A clinical and experimental study of fifty men of college age*. New York: Oxford University Press.

Musek, J. (2007). A general factor of personality: Evidence for the big one in the five-factor model. *Journal of Research in Personality, 41,* 1213-1233.

Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén and Muthén.

Nagengast, G., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K., & Trautwein, U. (2011). Who took the "×" out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science, 22,* 1058-1066. doi: 10.1177/0956797611415540

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht, The Netherlands: Kluwer.

Neuenschwander, M. P., Vida, M., Garrett, J. L., & Eccles, J. S. (2007). Parents' expectations and students' achievement in two western nations. *International Journal of Behavioral Development, 31,* 594-602. doi: 10.1177/0165025407080589

Nichols, R. C., & Bilbro, W. C., Jr., (1966). The diagnosis of twin zygosity. *Acta Geneteticae Medicae et Gemellollogiae, 16,* 265-275.

Nimon, K. (2010). Regression commonality analysis: Demonstration of an SPSS solution. *Multiple Linear Regression Viewpoints, 36,* 10-17.

Nimon, K., Lewis, M., Kane, R., & Haynes, R. M. (2008). An R package to compute commonality coefficients in the multiple regression case: An introduction to the package and a practical example. *Behavior Research Methods, 40,* 457-466. doi: 10.3758/BRM.40.2.457

*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

Notestein, F. (1945). Population: The long view. In T. Schultz (ed.), *Food for the world* (pp. 36-57). Chicago, IL: University of Chicago Press.

Obschonka, M., Schmitt-Rodermund, E, Silbereisen, R. K., Gosling, S. D., Potter, J. (2013). The regional distribution and correlates of an entrepreneurship-prone personality profile in the United States, Germany, and the United Kingdom: A socioecological perspective. *Journal of Personality and Social Psychology, 105,* 104-122.

Oishi, S., & Graham, J. (2010). Social ecology: Lost and found in psychological science. *Perspectives on Psychological Science, 5,* 356-377.

Ones, D. S. & Viswesvaran, C. (1996). Bandwidth – fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior, 17,* 609-626.

Ones, D. S. & Viswesvaran, C. (2001a). Integrity tests and other criterion-focused occupational personality scales (COPS) used in personnel selection. *International Journal of Selection and Assessment, 9,* 31-39.

Ones, D. S. & Viswesvaran, C. (2001b). 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). Washington, D. C.: American Psychological Association.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology, 78,* 679-703.

Oppenheimer, V. K. (1988). A theory of marriage timing. *American Journal of Sociology, 94,* 563-591.

Ou, S. & Reynolds, A. J. (2008). Predictors of educational attainment in the Chicago Longitudinal Study. *School Psychology Quarterly, 23,* 199-229. doi: 10.1037/1045-3830.23.2.199

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

Paunonen, S. V. (1998). Hierarchical organization of personality and prediction of behavior. *Journal of Personality and Social Psychology, 74,* 538-556.

Paunonen, S. V. & Ashton, M. C. (2001a). Big five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology, 81,* 524-539.

Paunonen, S. V. & Ashton, M. C. (2001b). Big five predictors of academic achievement. *Journal of Research in Personality, 35,* 78-90.

Paunonen, S. V. & Jackson, D. N. (1996). The Jackson Personality Inventory and the Big Five Model of personality. *Journal of Research in Personality*, *30*, 42-59.

Paunonen, S. V., Haddock, G., Forsterling, F., & Keinonen, M. (2003). Broad versus narrow personality measures and the prediction of behavior across cultures. *European Journal of Personality, 17,* 413-433.

*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(1998090)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

Pilia, G., Chen, W.-M., Scuteri, A., Orrú, M., Albai, G., Lai, S.,… Schlessinger, D. (2006). Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genetics, 2,* 1207-1223.

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, 88,* 245-258. 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

Poropat, A. E. (2009). A meta-analysis of the Five-Factor Model of personality and academic performance. *Psychological Bulletin, 135,* 322-338. doi: 10.1037/a0014996

Potter, J. E., Schmertmann, C. P., & Cavenaghi, S. M. (2002). Fertility and development: Evidence from Brazil. *Demography, 39,* 739-761.

Purtell, K. M., & McLoyd, V. C. (2013). Parents' participation in a work-based anti-poverty program can enhance their children's future orientation: Understanding

pathways of influence. *Journal of Youth and Adolescence, 42,* 777-791. doi: 10.1007/s10964-012-9802-7

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.

Raine, A., Reynolds, C., Venables, P. H., Mednick, S. A. (2002). Stimulation seeking and intelligence: A prospective longitudinal study. *Journal of Personality and Social Psychology, 82,* 663-674. doi: 10.1037/0022-3514.82.4.663

*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

Réale, D., Martin, J., Coltman, D. W., Poissant, J., & Festa-Bianchet, M. (2009). Male personality, life-history strategies, and reproductive success in a promiscuous mammal. *Journal of Evolutionary Biology, 22,* 1599-1607.

Reiss, D. (2003). Child effects on family systems: Behavioral genetic strategies. In A. C. Crouter & A. Booth (Eds.), *Children's influence on family dynamics: The neglected side of family relationships* (pp. 3-25). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Rentfrow, P. J., Gosling, S. D., & Potter, J. (2008). A theory of the emergence, persistence, and expression of geographic variation in psychological characteristics. *Perspectives on Psychological Science, 3,* 339-369.

Rentfrow, P. J., Gosling, S. D., Jokela, M., Stillwell, D. J., Kosinski, M., & Potter, J. (2013). Divided we stand: Three psychological regions of the United States and their political, economic, social, and health correlates. *Journal of Personality and Social Psychology, 105,* 996-1012. doi: 10.1037/a0034434

Rentfrow, P. J., Jost, J. T., Gosling, S. D., & Potter, J. (2009). Statewide differences in personality predict voting patterns in 1996-2004 presidential elections. In J. T. Jost, A. C. Kay, and H. Thorisdottir (Eds.), *Social and Psychological Bases of Ideology and System Justification* (pp. 314-347). Oxford University Press.

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

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., & Bogg, T. (2004). A longitudinal study of the relationships between conscientiousness and the social-environmental factors and substance-use behaviors that influence health. *Journal of Personality, 72,* 325-354.

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., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2,* 313-345. doi: 10.1111/j.1745-6916.2007.00047.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.

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: Lawrence Erlbaum Associates.

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: 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

Rosnow, R. L., Rosenthal, R., & Rubin, D. B. (2000). Contrasts and correlations in effect size estimation. *Psychological Science, 11,* 446-453.

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

Rowell, R. K. (1996). Partitioning predicted variance into constituent parts: How to conduct regression commonality analysis. *Advances in Social Science Methodology, 4,* 33-43.

Rushton, J. P. & Irwing, P. (2008). A general factor of personality (GFP) from two meta-analyses of the Big Five: Digman (1997) and Mount, Barrick, Scullen, and Rounds (2005). *Personality and Individual Differences, 45,* 679-683.

Rushton, J. P. & Irwing, P. (2009). A general factor of personality in 16 sets of the big five, the Guilford – Zimmerman Temperament Survey, the California Psychological Inventory, and the Temperament and Character Inventory. *Personality and Individual Differences, 47,* 558-564.

Rushton, J. P. & Irwing, P. (2011). The general factor of personality: Normal and abnormal. In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbook of individual differences* (pp. 132-161). West Sussex, UK: Wiley-Blackwell.

Rushton, J. P., Bons, T. A., & Hur, Y. (2008). The genetics and evolution of the general factor of personality. *Journal of Research in Personality, 42,* 1173-1185.

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin, 94,* 18-38.

Sameroff, A. (2009). *The transactional model of development: How children and contexts shape each other.* Washington, DC: American Psychological Association.

Santelli, J. S., & Melnikas, A. J. (2010). Teen fertility in transition: Recent and historical trends in the United States. *Annual Review of Public Health, 31,* 371-383. doi: 10.1146/annurev.publhealth.29.020907.090830

Saroglou, V. (2002). Religion and the five factors of personality: A meta-analytic review. *Personality and Individual Differences, 1,* 15-25.

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In Heijmans, R. D. H., Pollock, D. S. G. & Satorra, A. (Eds.), *Innovations in multivariate statistical analysis*. *A festschrift for Heinz Heudecker* (p. 233-247). London: Kluwer Academic Publishers.

Saucier, G. (1992). Benchmarks integrating affective and interpersonal circles with the Big Five personality factors. *Journal of Personality and Social Psychology*, *62*, 1025-1035.

*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: Oxford University Press.

Scarr, S. & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54,* 424-435. DOI: 10.2307/1129703

Scarr, S. (1997). Behavior genetic and socialization theories of intelligence: Truce and reconciliation. In R. J. Sternberg and E. L. Grigorenko (Eds.), *Intelligence: Heredity and environment* (pp. 3-41). New York: Cambridge University Press.

Schmertmann, C. P. (2003). A system of model fertility schedules with graphically intuitive parameters. *Demographic Research, 9,* 81-110.

Schmertmann, C. P. (2012). Calibrated spline estimation of detailed fertility schedules from abridged data. MPIDR Working Paper 2012-022.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8,* 350-353.

Schneider, B., & Stevenson, D. (1999). *The ambitious generation: America's teenagers, motivated but directionless.* New Haven, CT: Yale University Press.

Schneider, B., Keesler, V., & Morlock, L. (2010). The effects of family on children's learning and socialization. In H. Dumont, D. Istance, & F. Benavides (Eds.), *The nature of learning: Using research to inspire practice* (pp. 251-284). Paris: OECD.

Schofield, T. J., Martin, M. J., Conger, K. J., Neppl, T. M., Donnellan, M. B., & Conger, R. D. (2011). Intergenerational transmission of adaptive functioning: A test of the interactionist model of SES and human development. *Child Development, 82,* 33-47. doi: 10.1111/j.1467-8624.2010.01539.x

Sewell, W. H. & Hauser, R. M. (1972). Causes and consequences of higher education: Models of the status attainment process. *American Journal of Agricultural Economics, 54,* 851-861. doi: 10.2307/1239228

Sewell, W. H. & Hauser, R. M. (1980). The Wisconsin Longitudinal Study of social and psychological factors in aspirations and achievements. *Research in Sociology of Education and Socialization, 1,* 59-99.

Sewell, W. H. & Shah, V. P. (1967). Socioeconomic status, intelligence, and the attainment of higher education. *Sociology of Education, 40,* 1-23. doi: 10.2307/2112184

Seyfarth, R. M., Silk, J. B., & Cheney, D. L. (2012). Variation in personality and fitness in wild female baboons. *Proceedings of the National Academy of Sciences, 109,* 16980-16985.

Shedlin, M. G., & Hollerbach, P. E. (1981). Modern and traditional fertility regulation in a Mexican community: The process of decision making. *Studies in Family Planning, 12,* 278-296.

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: 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

Sibley, C. G., & Duckitt, J. (2010). The personality bases of ideology: A one-year longitudinal study. *The Journal of Social Psychology, 150,* 540-559.

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

Simpkins, S. D., Fredricks, J. A., & Eccles, J. S. (2012). Charting the Eccles' expectancy-value model from mothers' beliefs in childhood to youths' activities in adolescence. *Developmental Psychology*, *48,* 1019-1032. doi: 10.1037/a0027468

Skirbekk, V., & Blekesaune, M. (2013, August 5). Personality traits increasing important for male fertility: Evidence from Norway. *European Journal of Personality*. Advanced online publication. doi: 10.1002/per.1936

Smith, H. J., Sheikh, H. I., Dyson, M. W., Olino, T. M., Laptook, 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

Snow, K., Derecho, A., Wheeless, S., Lennon, J., Rosen, J., Rogers, J., . . . Einaudi, P. (2009). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), kindergarten 2006 and 2007 data file user's manual (2010-010)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics, Institute of Education Sciences.

Sonfield, A., & Gold, R. B. (2012). *Public funding for family planning, sterilization and abortion services, FY 1980-2010.* New York: Guttmacher Institute.

Soto, C. J., & John, O. P. (2009). Using the California Psychological Inventory to assess the Big Five personality domains: A hierarchical approach. *Journal of Research in Personality*, *43*, 25-38.

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 Pscyhology, 94,* 718-737. doi: 10.1037/0022-3514.94.4.718

Spearman, C. (1927). *The abilities of man*. London: Macmillan.

*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

Steiger, J. H. (1989). *EZPath: A supplementary module for SYSTAT and SYGRAPH*. Evanston, IL: SYSTAT.

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence, 35,* 401-426. doi: 10.1016/j.intell.2006.09.004

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

Sullins, D. P. (2013). Institutional selection for conformity: The case of U.S. Catholic priests. *Sociology of Religion, 74,* 56-58. doi: 10.1093/socrel/srs053

Surkyn, J. & Lesthaeghe, R. (2004). Values orientations and the second demographic transition (SDT) in Northern, Western and Southern Europe: An update. *Demographic Research, S3* (3). Rostock: Max Planck Institute for Demographic Research.

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., Löckenhoff, C. E., Crum, R. M., Bienvenu, O. J., & Costa, P. T., Jr. (2008). Five-factor model personality profiles of drug users. *BMC Psychiatry, 8,* art. no. 22.

Terracciano, A., Löckenhoff, C. E., Zonderman, A. B., Ferrucci, L., & Costa, P. T., Jr. (2008). Personality predictors of longevity: Activity, emotional stability, and conscientiousness. *Psychosomatic Medicine, 70,* 621-627.

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., …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

Terracciano, A., Sutin, A. R., McCrae, R. R., Deiana, B., Ferrucci, L., Schlessinger, D., Uda, M., & Costa, P. T., Jr. (2009). Facets of personality linked to underweight and overweight. *Psychosomatic Medicine, 71,* 682-689.

Theophrastus (1870). *The characters of Theophrastus* (R. C. Jebb, Trans.). London: Macmillan.

Thomson, E. (2002). Motherhood, fatherhood, and family values. In R. Lesthaeghe (Ed.), *Meaning and choice: Value orientations and life course decisions* (pp. 251-272). The Hague: NIDI/CBGS Publications No. 37.

Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., and Najarian, M. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Combined User's Manual for the ECLS-K Eighth-Grade and K–8 Full Sample Data Files and Electronic Codebooks (NCES 2009–004)*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Tucker-Drob, E. M. (2012). Presechools reduce early academic-achievement gaps: A longitudinal twin approach. *Psychological Science, 23,* 310-319. doi: 10.1177/0956797611426728

Tucker-Drob, E. M., Briley, D. A., & Harden, K. P. (2013). Genetic and environmental influences on cognition across development and context. *Current Directions in Psychological Science*, 22, 349-355. doi: 10.1177/0963721413485087

Tucker-Drob, E. M. & Harden, K. P. (2012a). Early cognitive development and parental cognitive stimulation: Evidence for reciprocal gene – environment transactions. *Developmental Science, 15,* 250-259. doi: 10.1111/j.1467-7687.2011.01121.x

Tucker-Drob, E. M., & Harden, K. P. (2012b). Learning motivation mediates gene-by-socioeconomic status interaction on early mathematics achievement. *Learning and Individual Differences, 22,* 250-259. doi:10.1016/j.lindif.2011.11.015

Tucker-Drob, E. M., & Harden, K. P. (2013). Gene-by-preschool interaction on the development of early externalizing problems. *Journal of Child Psychology and Psychiatry, 54,* 77-85. doi:10.1111/j.1469-7610.2012.02578.x

Tucker-Drob, E. M., Rhemtulla, M., Harden, K. P., Turkheimer, E., & Fask, D. (2011). Emergence of a gene × socioeconomic status interaction on infant mental ability between 10 months and 2 years. *Psychological Science, 22,* 125-133. doi: 10.1177/0956797610392926

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

Turkheimer, E., D'Onofrio, B. M., Maes, H. H., & Eaves, L. J. (2005). Analysis and interpretation of twin studies including measures of the shared environment. *Child Development, 76,* 1217-1233. doi: 10.1111/j.1467-8624.2005.00845.x-i1

Tynkkynen, L., Tolvanen, A., & Salmela-Aro, K. (2012). Trajectories of educational expectations from adolescence to young adulthood in Finland. *Developmental Psychology, 48,* 1674-1685. doi: 10.1037/a0027245

U.S. Census Bureau (2010). *American Community Survey*. Available at www.census.gov/acs/www/

U.S. Census Bureau (2010). *TIGER Products*. Available at www.census.gov/geo/maps-data/data/tiger.html

Udry, J. R. (2003). How to spin straw into gold. In A. C. Crouter & A. Booth (Eds.), *Children's influence on family dynamics: The neglected side of family relationships* (pp. 3-25). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Upadhyay, U. D., Weitz, T. A., Jones, R. K., Barar, R. E., & Foster, D. G. (2013, August 15). Denial of abortion because of provider gestational age limits in the United States. *American Journal of Public Health*. Advanced online publication. doi: 10.2105/AJPH.2013.301378

van de Kaa, D. J. (1987). Europe's second demographic transition. *Population Bulletin, 42,* 1-59.

Van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The general factor of personality: A meta-analysis of big five intercorrelations and a criterion-related validity study. *Journal of Research in Personality, 44,* 315-327.

Vecchione, M., Alessandri, G., Barbaranelli, C., & Caprara, G. (2011). Higher-order factors of the big five and basic values: Empirical and theoretical relations. *British Journal of Psychology, 102,* 478-498.

Ventura, S. J. (2009). Changing patterns of nonmarital childbearing in the United States. *NCHS Data Brief,* No. 18.

Verweij, K. J. H., Zietsch, B. P., Medland, S. E., Gordon, S. D., Benyamin, 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

Veselka, L., Schermer, J. A., Petrides, K. V., & Vernon, P. A. (2009). Evidence for a heritable general factor of personality in two studies. *Twin Research and Human Genetics, 12,* 254-260.

*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,* e102. doi:10.1038/tp.2012.27

Wang, M.-T. (2012). Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs, and interests. *Developmental Psychology, 48,* 1643-1657. doi: 10.1037/a0027247

Wang, M.-T., & Sheikh-Khalil, S. (2013). Does parental involvement matter for student achievement and mental health in high school? *Child Development*. Online first publication. doi: 10.1111/cdev.12153

Watkins, S. C. (1990). From local to national communities: The transformation of demographic regimes in Western Europe, 1870-1960. *Population and Development Review, 16,* 241-272.

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 linage scans of anxiety-related phenotypes. *European Journal of Human Genetics, 20,* 1078-1084. doi:10.1038/ejhg.2012.47

Westoff, C. F., & Jones, E. F. (1979). The end of "Catholic" fertility. *Demography, 16,* 209-217.

Wink, P., Ciciolla, L., Dillon, M., & Tracy, A. (2007). Religiousness, spiritual seeking, and personality: Findings from a longitudinal study. *Journal of Personality, 75,* 1051-1070.

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

Woodard, C. (2011). *American nations: A history of the eleven rival regional cultures of North America*. Penguin: New York, NY.

Yamagata, S., Suzuki, A., Ando, J., Ono, Y., Kijima, N., Yoshimura, K., Ostendorf, F., Angleitner, A., Riemann, R., Spinath, F. M., Livesley, W. J., & Jang, K. L.

(2006). Is the genetic structure of human personality universal? A cross-cultural twin study from North America, Europe, and Asia. *Journal of Personality and Social Psychology*, *90*, 987-998..

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics, 88,* 76-82. doi: 10.1016/j.ajhg.2010.11.011

*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

Zhang, Y., Haddad, E., Torres, B., & Chen, C. (2011). The reciprocal relationships among parents' expectations, adolescents' expectations, and adolescents' achievement: A two-wave longitudinal analysis of the NELS data. *Journal of Youth and Adolescence, 40,* 479-489. doi: 10.1007/s10964-010-9568-8

Zhao, H-K., Chang, C-W., Geng, N., Gao, L., Wang, J., … & Gao, G-D. (2012). Associations between personality changes and nucleus accumbens ablation in opioid addicts. *Acta Pharmacologica Sinica, 33,* 855-593. doi: 10.1038/aps.2012.10

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29,* 663-676. doi: 10.3102/00028312029003663

Zimmerman, 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

Zuckerman, M. (1992). What is a basic factor and which factors are basic? Turtles all the way down. *Personality and Individual Differences, 13,* 675-681.

Zuckerman, M., Kuhlman, D. M., Joireman, J., Teta, P. & Kraft, M. (1993). A comparison of three structural models for personality: The Big Three, the Big Five, and the Alternative Five. *Journal of Personality and Social Psychology, 65,* 757-768.