The Search is On for Missing Heritability – and Missing "Environmentality" – of Personality: A Developmental Approach

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Lindsay K. Matteson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Matt McGue

January 2016

Abstract

Twin and adoption research has consistently found evidence that genetic and unique environmental factors each account for about half of the observed individual differences in personality traits. However, studies aimed at finding associated genetic polymorphisms have had limited success. The inability to link genetic variants to personality phenotypes has become commonly known as the "missing heritability" problem (Manolio et al., 2009). Similarly, Bleidorn, Kandler, and Caspi (2014) recently named the difficulty in identifying specific environmental factors as causal influences on personality phenotypes the problem of "missing environment." One factor that may contribute to the difficulty in uncovering specific influences on personality is that the phenotype changes over time. In study 1, I performed bivariate Cholesky decomposition of personality traits across ages 14 and 29; genetic influences on personality were largely stable over this developmental period, but not entirely so (mean r_g = .75). Next, I examined the extent to which our assumptions about the additive nature of genetic influence on personality are true by performing univariate Genome-Wide Complex Trait Analysis (GCTA) separately at ages 17, 24, and 29. Restricting GCTA to specific developmental groups did not produce estimates of SNP heritability that were consistently different from those based on an adult sample. Significant SNP heritabilities were observed for the MPQ primary scales aggression, harm avoidance, and traditionalism as well as for the super-factor constraint; results suggested that SNPs could account for about half of the twin-estimated heritability for these traits (~20%). In study 2, I investigated the effects of getting married and having children on personality development across ages 17 and 29; after demonstrating a small

but significant relationship between these role transitions and MPQ super-factors constraint and negative emotionality, I followed up with a co-twin control analysis to determine whether such effects were causal. There were no significant differences in personality change between identical twins discordant for these role transitions, suggesting that associations could be attributed to familial factors underlying both personality and initiation of adult social roles.

Table of Contents

List of Tables……iv

List of Figures……...vi

Chapter 1……..1

Chapter 2……..15

Chapter 3……..52

Chapter 4……..94

Bibliography……104

List of Tables

Table 1. Description of primary scales of the Multidimensional Personality Questionnaire (MPQ)…….40

Table 2. Number of genotyped, Caucasian individuals and complete twin pairs with phenotypes at each target age…….41

Table 3. Mean (SD) MPQ scale scores and N by age and sex……...42

Table 4. Rank order stability (Pearson r and CIs) of all traits across assessments….43

Table 5. Twin correlations and standardized parameter estimates (CIs) from the univariate Cholesky decomposition model…….44

Table 6. Genetic and environmental correlations (CIs) from the bivariate Cholesky decomposition model……..46

Table 7. Univariate SNP heritability estimates (SE) and N from GCTA at different ages…….47

Table 8. Number of participants with personality data across marriage group, time, sex, and zygosity……77

Table 9. Number of participants with personality data across parenthood group, time, sex, and zygosity……78

Table 10. Mean (SD) personality trait T-scores by sex and age…….79

Table 11. Tests of fixed effects from linear mixed models considering marriage effects at the individual level……..80

Table 12. Parameter estimates from linear mixed models considering marriage effects at the individual level……...81

Table 13. Tests of fixed effects from linear mixed models considering parenthood effects at the individual level……82

Table 14. Parameter estimates from linear mixed models considering parenthood effects at the individual level……83

Table 15. Tests of fixed effects from linear mixed models considering marriage effects at the pair level……..84

Table 16. Parameter estimates from linear mixed models considering marriage effects at the pair level……..85

Table 17. Tests of fixed effects from linear mixed models considering parenthood effects at the pair level……..86

Table 18. Parameter estimates from linear mixed models considering parenthood effects at the pair level……..87

List of Figures

Figure 1. Change in T-scores over time for 3 super-factors……48

Figure 2. Additive effects of all traits by age in the offspring sample, including the composite phenotype……...49

Figure 3. SNP heritability of all traits across age in the offspring sample, including the composite phenotype……...50

Figure 4. Average heritability estimates, GCTA vs. twin model……51

Figure 5. Mean differences in constraint at ages 17 and 29 across pair-level marriage groups, separated by zygosity…….88

Figure 6. Mean differences in negative emotionality at ages 17 and 29 across pair-level marriage groups, separated by zygosity……..89

Figure 7. Mean differences in positive emotionality at ages 17 and 29 across pair-level marriage groups, separated by zygosity……..90

Figure 8. Mean differences in constraint at ages 17 and 29 across pair-level parent groups, separated by zygosity…….91

Figure 9. Mean differences in negative emotionality at ages 17 and 29 across pair-level parent groups, separated by zygosity……..92

Figure 10. Mean differences in positive emotionality at ages 17 and 29 across pair-level parent groups, separated by zygosity……..93

Chapter 1. Introduction

Why are we the way we are? It seems a simple enough question, but decades of research have not been able to nail it down to a few determining factors. When we ask this question, the characteristics in which we are most interested tend to be related to what we define as personality: the enduring and pervasive ways we think, feel, and behave. Beyond simply providing an opportunity for introspection or lively dinner conversation, knowing how personality develops may be of interest for several practical reasons. Traits have been shown to predict who will achieve most in work and education as well as who will get divorced (Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007), who will develop psychopathologies such as depression (Terracciano et al., 2010), and even who will live the longest (Kern & Friedman, 2008).

Because individuals are unique and behavior is undoubtedly multiply determined, it is difficult to disentangle relevant influences on development and observe causal effects, especially considering the large extent to which factors are often correlated. One promising approach to this problem is to use behavior genetic methodology, which is founded on the fact that family members are to some degree phenotypically similar and to some degree genotypically similar. Often, it is assumed that familial similarity is due to a shared environment, which would include the influence of factors such as socioeconomic status, geographical region, parenting styles, and shared family experiences. However, that is not the only possibility. Family members also share genetic material. Traditional

behavior genetic methodology relies on so-called "natural experiments" to isolate the different sources of variation among people in a particular variable, such as personality.

Twins are an experiment of nature; in particular, monozygotic (MZ) twins share an identical genetic code as well as their prenatal environment, both of which are possible influences on individual differences. When MZ twins are raised in the same home, the extent to which they are different presumably depends entirely on their own unique environments—those experiences, relationships, identities, etc. that they do not share with their twin. Their similarity, on the other hand, may be due either to shared genes or environment. One way to address this is to compare MZ similarity to that of other relatives who are raised together, yet do not share as much DNA. Dizygotic (DZ) twins raised together are the best comparison, as they too share a prenatal environment and any other relevant factors specific to being a twin, but they are only half as genetically similar (across segregating DNA) as are MZ twins. If we observe that MZ twins are more similar in some characteristic than are DZ twins, we may infer that there is some genetic influence on that characteristic (this requires an assumption that, other than differences in genetic similarity, MZ and DZ twins share all phenotypically relevant factors to the same degree; see Felson, 2014 for a discussion of such assumptions). Adoption studies are an experiment of society. In this case, genetically similar individuals are not raised in the same environment. Of interest is the degree of similarity between individuals and their adoptive parents and siblings; if they are substantially similar, then that similarity must be due to their shared environment because they do not have genetic material in common.

Numerous studies have supplied consistent evidence of genetic influences on personality. Overall, it appears that about 40% of the differences among people in personality traits can be attributed to genetic variability while the other 60% can be accounted for by environment (Vukasovic & Bratko, 2015). What has been most striking is the finding that relevant environmental factors are not those that are shared by family members but instead almost exclusively those that are unique to individuals. This implies that, even though siblings may share many aspects of their childhood and adolescent environments, those objectively shared factors are interpreted subjectively—uniquely—and are more likely to make those siblings different than to make them similar.

Understanding the *type* of influences that are relevant to personality is the first step in identifying causal determinants of personality variation, but it is not the end of the story. Advances in technology have allowed more precise and efficient investigation of genetic variants and their impact on phenotypes, and more sophisticated statistical and measurement techniques have been developed which have allowed for better modeling of phenotypes and their relationships. However, it has still been difficult to uncover the specific genetic or environmental influences relevant to personality. For example, studies aimed at finding associated genetic polymorphisms, such as genome-wide association studies (GWAS), have had limited success, and this inability to link genetic variants to personality phenotypes has become commonly known as the "missing heritability" problem (Manolio et al., 2009). Likewise, in a review of several studies, Turkheimer and Waldron (2000) found that those variables not objectively shared by siblings, including differential parenting, peer groups, sibling interactions, teacher relationships, and family

constellation, had very little effect on their phenotypic differences, including personality. This is surprising given the emphasis developmental researchers have placed on these kinds of variables. Bleidorn, Kandler, and Caspi (2014) dubbed this the "missing environment" problem. It is possible that we are mistaken in our conclusions about the type of influences we are looking for, but it is also possible that there is still a substantial amount of error or "noise" in our measurements. For example, while the average heritability estimate for personality traits has been estimated at .40 (Vukasovic & Bratko, 2015), the estimate is substantially higher (.60-.80) when personality is measured by multiple reporters (Riemann, Angleitner, & Strelau, 1997; Kandler, Riemann, Spinath, & Angleitner, 2010). One particular complication for personality phenotypes is the fact that even though personality is by nature a stable characteristic of a person, it does change over time to some extent.

The terms "stability" and "change" are ambiguous and seemingly contradictory. What do these terms really mean? How can personality be both consistent and flexible? It turns out we can describe the development of personality in multiple meaningful ways, each of which allows insight into a different aspect of this issue. *Differential stability* reflects the degree to which individual differences are maintained over time. That is, it refers to the consistency of rank-order in a population assessed on multiple occasions. It is measured by retest correlations that indicate, for example, how likely it is that the highest scoring individual within a group will remain the highest scorer at a later test administration, regardless of the change in actual personality scores. *Absolute change* informs us about how and when traits change on both a mean level and an individual

level. Population means of individual traits can be compared across time with analysis of variance (ANOVA). These means indicate the average amount and direction of change experienced by members of the population under study. These changes are often referred to as "normative" patterns of growth or maturity.

The pattern of rank-order stability observed across the life span suggests that personality traits are highly stable over time; however, even though stability increases across the life span, there remains a substantial amount of change even late in life when stability coefficients are highest. In their meta-analysis of differential stability, Roberts and Del Vecchio (2000) discovered that personality consistency increases with age, peaking between ages 50 and 70. Specifically, retest correlations averaged 0.43 in adolescence, grew to 0.60 in young adulthood, and peaked at 0.74 in old age (the authors corrected for varying time intervals across studies, and results are based on a mean interval of 6.7 years). Additionally, they found this pattern applied across all traits and across both sexes. More recently, Lucas and Donnellan (2011) analyzed a large ($N > 20,000$), nationally representative sample using latent variable modeling; this analysis has the advantage of correcting for measurement error. They found that personality stability continued to increase well beyond age 30, and in fact did not peak until about age 65. Caspi, Roberts, and Shiner (2005) referred to this as the *cumulative continuity principle* of personality, which implies that while there is a fair amount of change across the lifespan, personality tends to increase in stability rather than decrease as people age.

On the one hand, developmental change can be a nuisance when investigating the specific factors that contribute to personality, as samples often include adolescents as

well as adults of all ages; if there are age-related effects on the phenotype of interest, these effects may skew the results of studies examining the etiology of such phenotypes. Generally, age effects are included in behavioral genetic models, but often it is only to account for their moderating effects on the other variables of interest. On the other hand, change in personality is interesting in and of itself. There is substantial evidence that most people change in the same direction at around the same time. The pattern of normative changes that occur across young adulthood can be described by the *maturity principle*, which suggests that people grow in ways that allow them to function as productive members of society (Caspi et al., 2005). One of the most consistent observations of personality change is the increase in conscientiousness that occurs between adolescence and middle age; findings from a large meta-analysis (Roberts, Walton, & Viechtbauer, 2006) and also from recent, large, cross-sectional (Allemand, Zimprich, & Hendriks, 2008; Soto, John, Gosling, & Potter, 2011) and longitudinal studies (Bleidorn, Kandler, Riemann, Angleitner, & Spinath, 2009; Donnellan, Conger, & Burzette, 2007; Hopwood et al., 2011; Vaidya, Gray, Haig, Mroczek, & Watson, 2008) suggest an effect size between a half and a whole standard deviation. These same studies also suggest a significant increase in agreeableness as well as a decrease in neuroticism, though the sizes of these effects are more contentious than that of conscientiousness.

What is responsible for these trends? Could it be that people are genetically programmed to mature during certain periods of their lives, or is the fact that life circumstances often change during this period of life what drives personality development? Does genetic maturation prompt personality change which then allows

individuals to engage successfully in adult roles, or does taking on new life challenges inspire change? Two prominent theories have been proposed to answer these types of questions.

Five Factor Theory and intrinsic maturation

McCrae and colleagues have clearly stated their position on the causal processes responsible for personality development. According to their Five Factor Theory (FFT) of personality, traits are equivalent to temperaments such that they are endogenous dispositions independent of environmental influences (McCrae et al., 2000). They refer to these dispositions as *basic tendencies* that are expressed through *characteristic adaptations*, the concrete behaviors through which personality traits are inferred, such as skills, habits, beliefs, roles, identities, and relationships. This theory allows for the influence of environment on characteristic adaptations, but it emphasizes that the underlying dispositions driving behavior are based solely in biology and that environments merely channel the expression of such dispositions. McCrae and colleagues have based this conclusion on evidence of the stability and heritability of personality as well as structural similarity across cultures. Given that environments can change a great deal over a lifetime, they reason that personalities should change a great deal as well if environment is important to personality development. As previously discussed, however, personality traits are stable over time, especially in adulthood. Additionally, findings of moderate heritabilities of traits suggest a genetic basis of personality, and because the genetic code does not change over time, they argue that stability is likely based primarily on genetic influences. Finally, they propose that structural similarity across cultures

suggests that the core of personality is unaffected by culture (a major source of environmental influence).

An important tenet of FFT is that personality develops through intrinsic maturation, such that the extent of change that does occur over time is independent of life experiences and is instead due to genetically determined processes. McCrae and colleagues have supported their argument with cross-cultural comparisons of mean-level change over time. If personality development is highly influenced by the environment, then groups of people with very different histories should change in different ways. They argue that because individuals change in much the same ways both within cultures (as discussed above) and across cultures (e.g., McCrae et al., 2000), environmental factors do not significantly affect personality change; rather, they assert that a common biological maturation process is much more likely. But how can biological processes contribute to change if the genetic code never changes? The answer lies in the expression of genes. Throughout the life course, genes are variably "turned on" or "off," providing one mechanism by which major biological events can be genetically influenced. Personality traits may be strongly influenced by, for example, the physical changes associated with puberty; it is also possible that personality is biologically timed to develop in specific ways to coincide with puberty and the major life transitions to come.

Social Investment Theory

The fact that more mean-level change occurs during young adulthood than during other periods does not require the acceptance of intrinsic maturation as the sole process of change. An alternative explanation is that personality change is caused by significant life

changes, including both singular events (e.g., loss of a loved one) and lasting contextual changes (e.g., marriage). Social Investment Theory (SIT; Roberts, Wood, & Smith, 2005) emphasizes the plasticity of personality while at the same time acknowledging the great degree of stability (e.g., Caspi et al., 2005). It does not propose that biological processes do not matter; instead, it incorporates such processes into the understanding of personality development for reasons highlighted by McCrae et al. (2000).

This perspective has been well explicated by Roberts, Wood, and Caspi (2008). Specifically, they proposed that environmental factors and stable person characteristics (e.g., genetic dispositions) interact in ways that both stabilize and change personality over time. Outlined in this transactional model are seven processes by which personality remains consistent throughout life. These include both active and evocative processes; most notably, people seek out environments and select into roles consistent with their personalities ("attraction" and "selection," respectively). Individuals also may selectively focus on information relevant to their personalities, much like confirmation bias ("reactance"). In addition, people sometimes change their environments to suit themselves ("manipulation") or leave environments that require too much change ("attrition"). Most obviously, perhaps, is that people experience reactions from others that strengthen the tendencies leading them to evoke such reactions in the first place ("evocation"). Finally, Roberts and colleagues proposed that a sense of identity gained over the life course both reflects and facilitates these processes ("identity clarity"). All of these processes are examples of the *corresponsive principle*, which states that "the most

likely effect of life experience is to deepen the characteristics that lead people to those experiences in the first place" (p. 470, Caspi et al., 2005).

Change, from this point of view, is largely a product of taking on new social roles (e.g., husband, parent, employee, leader, etc.) over the life course. Each new role carries with it a specific set of expectations, demands, and contingencies that together provide a guide for how individuals should behave in that role; thus, behavior may be modified to align with the new role requirements. Doing so involves responding to rewards and punishments, accepting feedback from others, and modeling others' behavior. Mean-level changes in personality reflect the tendency for most individuals to adapt to systematic role changes that occur across the life span, such as becoming more conscientious as one enters the workforce or produces offspring. For example, Roberts, Caspi, and Moffitt (2001) suggested that the greater maturity of females in young adulthood may be due to the fact that they are more likely than their male counterparts to have children (because women reproduce with men who are, on average, two years older). Cross-cultural similarities can be explained by the fact that most cultures encourage the same universal role changes at about the same ages. Around the world (or at least the industrialized world), marriage, family, and work are of prime importance during young adulthood and middle age while old age is instead concerned with transitioning into retirement, coping with the loss of loved ones, and becoming a grandparent, among other things (Roberts et al., 2006). By this theory, then, individual differences in the rate of change could possibly be explained by forces that influence the onset of new roles, such as the incidence of

specific life events or other environmental factors, the development of other psychological processes, or biological maturation.

Which came first, the chicken or the egg?

The difference between these perspectives is essentially a question of the direction of causal arrows: Do personality changes drive changes in social roles, or does the pursuit of new roles compel personality to change? In order to infer that the latter is true, one must be able to either prospectively predict changes in personality based on social context or provide causal evidence of change resulting from interventions (Roberts et al., 2008). Alternatively, research which utilizes genetically informative samples can shed light on the mechanisms responsible for personality change. In fact, Caspi et al. (2005) provided four reasons why quantitative genetic studies are important for understanding personality development. First of all, these studies can disentangle genetic and environmental influences, unlike more traditional regression studies. Secondly, results from these studies have helped explain why children in the same family are so different; a major revelation resulting from these studies is that instead of making siblings more similar, environmental factors actually tend to make siblings more different. Thirdly, behavior genetic methodology can examine the direction of causation between variables. For example, genetically informative studies can help explain why environmental factors are themselves heritable. Finally, behavior genetic methods are ideal for investigating age-related etiological change. Such methods can, for instance, distinguish between influences on the onset of psychopathology and intensity or persistence of symptoms, which are time-relevant.

Given that behavior genetic methods hold such promise for delineating the factors underlying personality maturation, I made use of such methods to examine the extent to which each developmental theory could be supported. Additionally, by incorporating a developmental perspective into the investigation of personality etiology, I further refined the phenotypes of interest in the hunt for the so-called missing heritability and environmentality of personality traits.

In study 1, my first goal was to describe the personality changes that occur from adolescence through early adulthood in both general and specific traits measured by the Multidimensional Personality Questionnaire (MPQ; Tellegen & Waller, 2008). My second goal was to examine the extent to which genetic influences were stable from age 14 to age 29, as a high degree of stability would suggest that gene-finding efforts may not be substantially impeded by using mixed-age samples. To do this, I calculated genetic and environmental correlations across ages based on bivariate Cholesky decomposition of variance for different traits. My third goal was to determine whether the assumption of primarily additive genetic effects on personality is warranted; that is, I conducted Genome-Wide Complex Trait Analysis (GCTA; Yang et al., 2011) which is a method that estimates the influence of all common, genotyped single nucleotide polymorphisms (SNPs) on a phenotype at once. If estimates of SNP heritability do not approach heritability estimates from twin studies, one possible explanation is that personality-relevant genetic influences are not additive in nature. To reduce age-related phenotypic heterogeneity, I conducted GCTA at ages 17, 24, and 29 separately and compared the estimates of SNP heritability to those from an adult sample.

In study 2, I investigated the impact of major life transitions on personality maturation between ages 17 and 29. Specifically, I examined differences in personality change between those who had ever married or had ever become parents and those who did not experience those transitions by age 29 in linear mixed models (with random effects to account for the nesting of variables). Though past research has demonstrated substantial associations between entering into these new roles and personality change, such studies are typically unable to provide evidence of causality. To address this limitation, I conducted co-twin control analyses, again with mixed modeling. In these analyses, I compared personality change across ages 17 and 29 in MZ twins discordant for social role transitions. Because MZ twins share 100% of their genes and most of their pre- and post-natal environments, they provide the best comparison group for a test of causal effects. If MZ twins who experienced the transition are no different than their twins who did not experience such a transition, then it would be likely that shared familial factors explain the relationship between social role transitions and personality maturation.

Together, these studies provide complementary information regarding specific influences on personality during a particular developmental period which may inform theories of development. For instance, evidence of increasing stability of environmental factors would be consistent with SIT because it proposes that people act in ways that strengthen pre-existing personality dispositions, and that the increasing ability of individuals to choose their environments as they enter into adulthood promotes these stabilizing processes even more. At the same time, evidence of causal effects of marriage

and parenthood on personality maturation would also provide support for SIT as it suggests that personality change is a product of taking on new social roles.

Chapter 2. SNP influence on personality: Can a developmental approach aid our search for "Missing Heritability"?

Twin and adoption research has consistently found evidence that individual differences in personality traits are about 40% heritable (Vukasovic & Bratko, 2015). However, studies aimed at finding associated genetic polymorphisms, such as genome-wide association studies (GWAS), have had limited success. The inability to link genetic variants to complex phenotypes has become commonly known as the "missing heritability" problem (Manolio et al., 2009). There are many possible explanations for this problem, including the prospect that years of twin research have led us to inaccurate heritability estimates. After all, twin studies rely on assumptions that may not in reality be met. For instance, the Equal Environments Assumption (EEA) suggests that monozygotic and dizygotic twins do not differ in the degree of similarity they experience in their environments; if violated, heritability is overestimated. Additionally, twin studies are limited in their ability to estimate the influence of all relevant factors. Much past research is based on twins only, but this restricts modeling to include only additive or nonadditive genetic effects (and usually it is additive effects that are reported). If both nonadditive genetic factors and shared environmental factors contribute to variance in a phenotype but only additive effects are modeled, then overall heritability estimates may be inflated. Fortunately, we are not limited to twin studies to examine heritability; adoption studies and extended family studies also provide information regarding genetic effects. Indeed, studies do suggest significant nonadditive influences on personality (e.g.,

Eaves et al., 1999; Finkel & McGue, 1997; Matteson, McGue, & Iacono, 2013). Unfortunately, methods of gene discovery are generally able to identify variants with additive effects only.

Another possible explanation is that we are correct in our assumption of significant additive genetic effects, but it is variants that are rare in the population that significantly affect personality, not those that are common. Unfortunately, GWAS are based on data from microarrays that capture common variation but from which rare variants are poorly imputed (Zheng, Ladouceur, Greenwood, & Richards, 2012); this means that they may not be able to identify personality-relevant variants.

Furthermore, it is possible that there may be such a large number of variants affecting these complex phenotypes that any individual variant, such as a single nucleotide polymorphism (SNP), has an effect so small that extremely large sample sizes are required to detect them. In fact, most researchers now recognize that the effect of any specific variant must be much smaller than previously anticipated, explaining 0.5% or less of variance in traits, and this has led to the consolidation of data across multiple cohorts to increase sample size to the point that it is possible to detect such small effects. The first of these consortia for personality (de Moor et al., 2012) included over 17,000 individuals, thus increasing power to detect small effect sizes; however, the two hits discovered in this GWAS were not significant in the replication sample. More recently, the Genetics of Personality Consortium performed a meta-analysis of neuroticism GWAS including 30 cohorts and over 63,000 individuals (de Moor et al., 2015) which again increased the power to detect small effects. They found one genome-wide significant

SNP, but again, the association failed to replicate. However, an interesting aspect of this study was the use of polygenic scores to predict both neuroticism and Major Depressive Disorder (MDD). That is, they tested the predictive power of a weighted sum of the most significantly associated markers from the meta-analysis in a smaller sample. They were able to predict about 1% of the variance in both phenotypes with these scores, thus suggesting that consideration of the joint effects of SNPs may be a viable approach to identifying the underlying genetic structure of personality.

Given these possibilities, one way of testing the current assumption of a high degree of common additive SNP influence on personality (or any other complex phenotype) is to consider the effects of all common SNPs simultaneously; Yang et al. (2011a) developed a statistical method to do just this, called genome-wide complex trait analysis (GCTA). This method estimates the variance explained by all genotyped SNPs on a chromosome or across the entire genome. Generally, GCTA estimates SNP heritability by associating genetic relatedness among all individuals in the sample with phenotypic covariance. Studies on height and BMI (Yang et al., 2011b), and even cognitive ability (Davies et al., 2011) have successfully produced significant estimates of SNP effects with this method; in fact, Davies and colleagues were able to explain 40-51% of the phenotypic variance in cognitive ability with SNPs. A few GCTA studies of personality have been published, but estimates of SNP effects have been lower than expected. Verweij and colleagues (2012) reported total SNP effects ranging from .04 to .10 for TPQ traits harm avoidance, novelty seeking, reward dependence, and persistence; similarly, Vinkhuyzen et al. (2012) reported SNP heritabilities of .06 for neuroticism and

.12 for extraversion. Rietveld et al. (2013) reported a slightly larger effect for subjective well-being (.15), and Power and Pluess (2015) observed significant SNP heritabilities of .15 and .21 for Big Five traits neuroticism and openness to experience, respectively. Although these estimates are not zero, they are indeed much lower than the heritability estimates from twin studies. Recently, a subset of participants in the neuroticism meta-analysis previously mentioned (de Moor et al., 2015) was included in GCTA; an advantage of this study is that the neuroticism phenotype was obtained from a harmonization of data from different personality inventories using item response theory (van den Berg et al., 2014). It can be difficult to compare results across studies when different personality assessments are used, but harmonizing scores makes use of what is common across inventories for a better phenotype. However, even with this improvement, SNP effects were still smaller than expected (.15). This would suggest that the majority of missing heritability for personality cannot be explained by common SNPs.

Of note, however, is the fact that the samples in the previous studies consisted of adults varying widely in age. The failure of genetic association studies to uncover significant variants may not simply be an issue of sample size; an important consideration that has typically been ignored in molecular research thus far is development. Vrieze et al. (2012) pointed out that a major impediment to our gene-finding efforts is that phenotypes may be too etiologically heterogeneous and complex to uncover gene-behavior associations. We know that personality changes systematically as individuals progress through developmental periods (increases in emotional stability,

conscientiousness, agreeableness, and social dominance; e.g., Roberts, Walton, & Viechtbauer, 2006), and the Five Factor theory of personality as put forth by McCrae and Costa (2008) suggests that intrinsic maturation is responsible for these changes. If this is the case, then different genetic influences may be relevant at different ages, and thus even large meta-analyses of genetic association studies may fail to detect effects due to the developmentally heterogeneous samples they combine.

Until recently, it is has been unclear whether heritability changes throughout the entire span of adulthood, as some have found decreases with age (Pederson & Reynolds, 1998; Viken et al., 1994) while others have observed no change (Heiman et al., 2003; Loehlin & Martin, 2001). However, a recent meta-analysis (Briley & Tucker-Drob, 2014) reported a significant decrease in heritability across Big Five traits as well as a parallel increase in environmentality across the life span. Specifically, in their continuous model, they observed a large decrease in heritability from early life to adolescence and a steady decrease in heritability across adolescence and adulthood (estimates dropped from approximately .75 to .34 across the entire lifespan); they also observed a steady increase in true (corrected for measurement error) non-shared environmental influences from approximately .10 to .45. The most parsimonious non-continuous models suggested that the majority of the decrease in heritability occurred by age 15 and that the majority of the increase in non-shared effects occurred by age 30.

In addition to considering heritability, Briley and Tucker-Drob also examined whether the same genetic (and environmental) factors were important across the life span by considering the magnitude of genetic (and environmental) correlations (r_g and r_e)

across an average of 5.6 years; these correlations suggested that, although genetic factors are constant across time even early in life ($r_g \sim .8$), there is at least some change in the genetic factors relevant to personality from adolescence to young adulthood because genetic correlations did not reach unity until after age 30. Additionally, research has provided evidence of genetic influence on change in some traits (agreeableness and conscientiousness) as measured by slope in a longitudinal growth curve model (Bleidorn, Kandler, Riemann, Angleitner, & Spinath, 2009; Hopwood et al., 2011). If the genetic changes are significant, then results of genetic association studies based on adult samples may not be applicable to younger samples. Furthermore, age-heterogeneous samples may reduce the signal-to-noise ratio in studies of adults. This begs the question of whether developmental noise has contributed to the small amount of variance that has been explained in previous GCTA studies of personality. Additionally, it raises the question of whether common SNPs can explain personality variation in adolescents and young adults to the same degree that it does in older adults.

In the current study, I estimated the influence of additive genetic factors on personality traits both biometrically and molecularly, and I did so at different target ages spanning the developmental period of adolescence through young adulthood. I asked the following questions:

Q1: To what extent does personality change in my sample from adolescence through young adulthood? Much of the previous research on this topic has considered only domain-level changes in personality, and I aimed to extend that knowledge by

considering change in individual scales of the Multidimensional Personality Questionnaire (MPQ) from age 14 to 29 in both mean level and rank order stability.

Q2: If personality does change significantly from adolescence through young adulthood, to what extent do genetic influences differ over time? Does heritability change significantly from age 14 to 29? Are the genes important to personality variance at age 14 the same as those that are important at ages 17, 24, and 29? Although previous researchers (Blonigen et al., 2008 and Hopwood et al., 2011) have reported genetic and environmental correlations from Minnesota Twin and Family Study (MTFS) data, they only included participants from the older cohort; in the current study I included participants from both the younger and older cohorts, which allowed me to both increase my sample size and extend findings to age 14.

Q3: If there are age differences in the relevant genetic influences on personality as estimated from biometrical models, will results of GCTA reflect those differences? That is, are there major developmental differences in the influence of additive genetic variation on personality? Will estimates of SNP heritability differ at ages 14, 17, 24, and 29? Furthermore, how might those estimates differ from those of an entirely adult sample? Additionally, if we consider only personality variance which is consistent across ages, will our ability to detect SNP influence increase?

Method

Participants

The current study included two samples, a twin offspring sample and an adult (parent) sample. Offspring came from the Minnesota Twin and Family Study (MTFS;

Iacono & McGue, 2002) only, while adults were drawn from the MTFS as well as the Sibling Interaction and Behavior Study (SIBS; McGue et al., 2007). The MTFS is an ongoing longitudinal study of reared-together, same-sex twins (N = 3779). Its primary focus is to identify the genetic and environmental bases of substance abuse and related psychopathology. The SIBS is a longitudinal study of different kinds of families, including those with 2 biological children, 1 adoptive child and 1 biological child, or 2 adoptive children (N = 617 families). Its focus is to identify how family environment impacts development. Exclusion criteria for both samples included living more than a day's drive from Minneapolis and presence of any mental or physical handicap that would prevent completion of the assessment.

Twin offspring. Twins were recruited at approximately ages 11 (younger cohort) or 17 (older cohort) from publicly available Minnesota birth certificates. Participants were followed up approximately every three to four years through their 20s, and participation rates were generally greater than 90% at each of the follow-ups. The MTFS sample was representative of the demographics of Minnesota for the birth cohorts sampled; as such, twins were mostly Caucasian (over 95%). Zygosity was determined by three different methods including staff opinion (based on similarity of face and ear shape and hair and eye color), the Physical Similarity Questionnaire (Peeters, Van Gestel, Vlietinck, Derom, & Derom, 1998), and an algorithm based on measurements of cephalic index (ratio of head width to length), fingerprint ridge counts, and ponderal index (a measure of leanness calculated as height (in)/$\sqrt[3]{\text{weight (lbs.)}}$). When measures did not agree, a serological sample was taken to determine zygosity. Phenotype data were

obtained from follow-ups targeted at ages 14 (younger cohort only) and 17, 24, and 29 (both cohorts) in order to cover the major developmental periods of adolescence and early adulthood. In total, data were available from at least one assessment for 3646 individual twins (52.7% female).

Adults. Parents from the MTFS and SIBS comprised the adult sample. Phenotype data were collected at the intake assessment in most cases. Phenotype data were available for 4647 of these adults (51.8% female).

Phenotypes

For participants at the age 17, 24, and 29 assessments as well as the parent assessment, personality was assessed with a 198-item version of the MPQ (Tellegen & Waller, 2008); participants at the age-14 assessment completed a shorter version of the MPQ (133-item Personality Booklet-Youth, Abbreviated (PBYA), developed specifically for the MTFS) that consisted of identical items but only 6 of 11 primary scales. Items are endorsed on a scale from *1 = Definitely True* to *4 = Definitely False*, and items are scored so that higher scores represent higher levels of the trait. The MPQ is a self-report personality inventory derived from factor analysis. It measures 11 primary personality traits and 3 higher-order factors. The higher-order factors represent the behavioral and emotional regulation that contributes to particular traits; they include positive emotionality (PEM), negative emotionality (NEM), and constraint (CON). The phenotypes of interest were all 11 primary scales and the 3 super-factors. Table 1 briefly describes the primary scales. Additionally, I created a composite phenotype ("COMP") for each scale that consisted of the average of all scores across assessments for every

participant for whom data were available for more than one assessment (the mean number of assessments included in the composite variable was approximately 2.75 across traits).

Genotyping

DNA samples were genotyped on Illumina 660W Quad array (Illumina, Inc., San Diego, CA), using NCBI genome build 36.2 as reference data. DNA samples and typed markers were both subjected to thorough quality-control procedures (see Miller et al., 2012); 7278 samples (including offspring and adults from both studies) survived quality control. Because only one twin in a pair of MZ twins needs to be typed, genome-wide SNP data were available for an effective total of 8405 individuals. EIGENSTRAT (Price et al., 2006) was used to extract the first 10 principal components, which were used as covariates in GCTA to control for population structure. After SNP quality control, 527829 SNPs remained useful for analysis.

Final genetic sample

Of the 3646 offspring with phenotype data from at least one personality assessment, genotypes were available for 3337. To prevent possible confounding due to population stratification, I excluded non-Caucasian participants from GCTA (132 individuals), leaving a total of 3205 individuals (52.2% female) for these analyses. Additionally, for GCTA, I only included genotyped adults of Caucasian descent (N = 3799, 54.3% female). In order to keep the analyses as comparable as possible, I included only genotyped Caucasian twins in the biometric analyses. Table 2 describes the sample sizes at each targeted age. Although participation rates were high at all follow-ups in the

overall study, personality data was not available for everyone at these follow-ups. Specifically, of those individuals who participated at age 17, 82.5% and 76.1% also participated at ages 24 and 29, respectively. However, I observed few significant differences in age 17 personality between those who participated at 24 and 29 and those who did not, and the significant differences I did observe were small (Cohen's *d*s ranged from .142 to .293).

Descriptive analyses

For each trait, I conducted two-factor, repeated-measures ANOVA for sex, assessment (target assessment age), and the sex-assessment age interaction to examine how personality changes over time at the mean level. Additionally, I calculated Pearson correlations across assessments to examine rank order stability.

Twin analyses

For each phenotype, I first calculated correlations between both MZ and DZ twins and then began biometric analyses. Model fitting with the classic twin design partitions phenotypic variance into additive genetic (A), shared environmental (C), and non-shared environmental (E) components. Within MZ twin pairs, both additive genetic and shared environmental effects are assumed to correlate 1.0, whereas within DZ twin pairs, shared environmental effects correlate 1.0 but additive genetic effects only correlate .5. Non-shared environmental influences are assumed to be uncorrelated and therefore only contribute to differences within pairs. I used Mx statistical software (Neale et al., 2002) to fit raw data to models with full-information maximum-likelihood estimation, which addresses missing data. Parameters are freely estimated from raw data by minimizing

minus twice the log-likelihood (-2LL), and the -2LL of more restricted models are compared to this with a likelihood ratio Chi square test of goodness of fit.

I conducted univariate analyses on all 14 phenotypes (11 primary scales + 3 super-factors) at ages 17, 24, and 29, on 6 phenotypes at age 14 (only 6 primary scales were assessed at that age), and on the composite scores. I first fit an ACE model to the sample of males and females combined; however, because ANOVA suggested significant sex differences in some scales (see Table 2), I also compared fit of a model where ACE estimates were allowed to differ across sexes to one where the standardized estimates a^2 and c^2 were constrained to be equal. A significant drop in model fit for the constrained model would suggest significant sex effects. Additionally, I conducted bivariate biometric analyses of personality across ages (14-17, 14-24, 14-29, 17-24, 17-29, and 24-29) to estimate the stability of genetic and environmental effects. I fit a bivariate Cholesky decomposition, but I present the results from a correlated factor solution because it is mathematically equivalent and easier to interpret (Loehlin, 1996).

Genome-wide Complex Trait Analysis (GCTA)

I used the standard GCTA software package (Yang et al., 2011) for all analyses. In this type of analysis, restricted maximum-likelihood estimation is used to separate a trait's variance into its genetic and residual components. First, it estimates a kinship coefficient for each pair of individuals in a sample (the genetic relatedness matrix, or, GRM) based on typed SNPs; this GRM can then be used to exclude one individual of each pair whose degree of relatedness is greater than a specified threshold in order to eliminate close relatives from the analysis. Exclusion is done because including close

relatives makes results less interpretable; it presents two possible confounds: 1) family members may share a common environment, and 2) family members may share genetic variation other than that captured by SNPs. Next, each SNP is entered as a random effect in a mixed linear model and heritability estimated using information on kinship differences. Because the GCTA estimate of heritability reflects only the effects of genotyped variants, results represent the lower-bound estimate of all additive genetic effects on the phenotype.

Cross-sectional age differences in the SNP contributions to personality were observed by performing univariate GCTA on all traits at the 17-, 24-, and 29-year old assessments (data were too limited for this analysis at age 14) and also for parents. The composite phenotypes were also included in these analyses. The kinship coefficient threshold for the GRM was set < .025 (fourth cousins), and I included sex and the 10 principal components mentioned earlier as covariates in all analyses.

Results

Descriptive Analyses

Table 3 displays the mean raw scores for each phenotype across assessment and sex as well as results from ANOVA. There were significant sex differences in every trait except absorption (AB), traditionalism (TR), and well being (WB). Partial η^2 for sex ranged from .000 to .202 and were largest for aggression (AG, .202, males higher), harm avoidance (HA, .184, females higher), and the constraint super-factor (CON, .096, females higher). Additionally, there were significant mean differences across assessments

in every trait except WB. Partial η^2 ranged from .002 to .351 and several were larger than .10, including AG (.351), negative emotionality (NEM, .266), alienation (AL, .231), and AB (.107) which all decreased in magnitude over time, and CON (.217), control (CN, .207), and HA (.161) which all increased in magnitude over time. Figure 1 illustrates change over time in the 3 super-factors represented by T-scores standardized to age 17. For several traits, I observed a significant interaction between sex and assessment (age); however, partial $\eta^2 < .02$ in all cases.

Within-individual cross-assessment correlations (rank order stability) are presented in Table 4; they ranged from .39 to .48 (M = .44) across the entire developmental period (ages 14-29) and all were statistically significant. Correlations increased with the age of the sample but decreased with the time interval; to illustrate: correlations were stronger between ages 14 and 17 (R = .54 to .62, M = .58), than between ages 14 and 24 (R = .37 to .51, M = .46), yet correlations between ages 14 and 17 were weaker than those between ages 24 and 29 (R = .70 to .82, M = .74) despite the longer interval.

Twin Correlations

Twin correlations are reported in Table 5. Correlations ranged from .32 to .65 for MZ twins and .05 to .42 for DZ twins. In all traits across all assessments, DZ correlations were smaller than MZ correlations. Considering the effect of time on twin similarity, the average (across traits) MZ correlation decreased slightly from age 14 (M = .51) to age 29 (M = .48) while the average DZ correlation increased from age 14 (M = .16) to age 29 (M = .24). Additionally, the correlations for the composite phenotypes (“COMP”) were

larger in most cases than at any individual assessment for both MZs (M = .59) and DZs (M = .29).

Univariate Twin Analyses

Results of univariate twin analyses are displayed in Table 5, and estimates of additive genetic effects on each phenotype are illustrated in Figure 2. All reported results are those from models in which standardized estimates of a^2 and c^2 were set equal between sexes, as this did not significantly decrease model fit in most cases, but did allow means to differ between males and females. The estimate of shared environmental influence was significantly different from zero in only one case, AG at age 14 (c^2 = .22), though there were additional non-zero estimates. In all phenotypes, the estimates of additive genetic influence were significantly different from zero and ranged from .31 to .59 across all traits and assessments while the estimates of non-shared environmental influence were all significantly different from zero and ranged from .34 to .69. The average a^2 estimate across traits decreased slightly from age 14 (M = .47) to age 29 (M = .45). Additionally, a^2 estimates for the composite phenotypes ranged from .44 to .66 (M = .54) and were in nearly all cases larger than estimates from individual time points while e^2 estimates for the composite phenotypes ranged from .29 to .56 (M = .42) and were in nearly all cases smaller than estimates from individual time points.

Bivariate Twin Analyses

Genetic and environmental correlations across assessments from the bivariate Cholesky models for each trait are presented in Table 6. As with the univariate biometric analyses, results are from the model where variance components were constrained across

males and females but means were allowed to vary. However, because estimates of c^2 were not significant in univariate analyses, bivariate analyses were based on an AE model. Considerable stability was observed in both environmental and genetic factors across all ages (all correlations were significantly different from zero). Genetic correlations were larger than environmental correlations in every phenotype, though only a few reached unity. Also, both genetic and environmental stability increased with age and with smaller intervals between assessments. The best illustration of these facts is observed if we consider the correlations between ages 14 and 17 versus those between longer intervals (such as 14-29) and those at later ages (such as 24-29). Specifically, from 14 to 17, the average genetic correlation was .83 (*R* = .67–1.00) and the average environmental correlation was .37 (*R* = .33-.41), both of which are larger than those from 14 to 29 when average r_g was .75 (*R* = .45-.93) and average r_e was .28 (*R* = .24-.34); at the same time, from 24 to 29, the average genetic correlation was .97 (*R* = .91-1.00) and the average environmental correlation was .54 (*R* = .43-.64), which are the highest correlations observed between any interval.

GCTA

Table 7 presents the results of univariate GCTA, and the estimates of SNP-based heritability (h^2_{SNP}) are displayed in Figure 3. In this relatively small developmental sample, standard errors were large in analysis of all phenotypes. Estimates of h^2_{SNP} varied considerably across assessments, ranging from 0 to .60, with only 2 significant estimates (SR age 24 = .60 and TR age 17 = .47). Composite phenotype estimates were non-zero for 8 traits, ranging from .03 to .35 (*M* = .12), but none were statistically significant. In

the sample of adults (parent generation), h^2_{SNP} ranged from 0 to .31 (M = .11) and standard errors were considerably smaller than those from the offspring sample. Estimates were largest (and statistically significant) for TR (.23), the super-factor CON (.23), AG (.22), and HA (.20).

Compared to estimates of additive genetic variance from the twin analyses, h^2_{SNP} estimates were much smaller. Figure 4 illustrates this point and shows that the average h^2_{SNP} estimate decreased slightly from age 17 to 29 (from .18 to .13) while the average twin estimates of a^2 changed only marginally.

Discussion

The current study is the first to consider development in the search for the "missing heritability" of personality. I examined phenotypic, genotypic, and environmental stability and change in both higher-order traits (positive emotionality, negative emotionality, and constraint) and more specific traits (11 primary scales of the MPQ) from adolescence through young adulthood.

First, I observed significant sex differences in most personality traits, but these effects were generally small (mean η^2 = .05). To accommodate, means were allowed to differ across sex in biometrical models and sex was included as a covariate in GCTA. Importantly, I observed that the mean levels of some personality traits changed from age 14 to 29 while others did not; negative emotionality decreased across this period while constraint increased. Positive emotionality did not change significantly. Like their corresponding super-factors, I observed significant decreases in aggression, alienation, and stress reaction (though the latter first increased from early to late adolescence before

decreasing into adulthood) and significant increases in control and harm avoidance (though the latter first decreased in males from age 14 to 17). Compared to sex effects, the effects of time on NEM- and CON-related traits were either comparable or much larger. However, the small effects that I observed for PEM-related primary traits were more mixed such that there were increases in achievement for both sexes and in social potency for men (and a decrease for women) but no differences across time for social closeness or well being. It is no surprise that these results are similar to those of Hopwood et al. (2011), as these samples overlap, but they are also corroborated by meta-analysis of Big Five traits (Roberts et al., 2006) as well as more recent longitudinal studies (e.g., Bleidorn et al., 2009) in which the authors reported significant increases in conscientiousness, agreeableness, and social dominance as well as decreases in neuroticism during this developmental period.

Second, I observed a substantial amount of stability over the entire developmental period. Phenotypic correlations were consistently larger than .4 across the 15 year span. As has been found in previous studies, correlations generally decreased with larger time intervals (Briley & Tucker-Drob, 2014; Fraley & Roberts, 2005) but increased with the age of assessment (Briley & Tucker-Drob, 2014; Ferguson, 2010; Roberts & DelVecchio, 2000). For instance, correlations between ages 24 and 29 were much larger than any other intervals (all *r*s > .7) even though it's not the shortest interval between assessments. Our finding that the rank order of individuals grew increasingly more stable reflects the cumulative continuity principle of personality development (Caspi et al., 2005). Still, I did observe substantial rank order change from adolescence to young adulthood.

Interestingly, the degree of rank order stability did not differ across traits, even though changes in mean level did differ across traits.

In the current study, although I corroborated previous findings of significant heritability at individual time points, I observed only a very slight decrease in additive genetic effects on personality traits over time, much smaller than suggested by meta-analysis (Briley & Tucker-Drob, 2014). However, our estimates of additive effects at age 29 were very similar to those at the same age in the meta-analysis.

Overall, the results of our bivariate biometric analyses coincided with the results of previous research (e.g., Briley & Tucker-Drob, 2014). Such analyses suggested high stability in genetic factors over time, as all genetic correlations were greater than .6 (with only 1 exception). Overall I observed an increase in genetic stability as participants aged; estimates of genetic correlations between ages 24 and 29 reached unity in many cases (*R* = .91-1.00) whereas they were smaller between all other ages. This suggests that developmentally relevant genetic factors stabilize by about age 30. On the other hand, I did observe evidence of change in genetic influences on personality from adolescence to adulthood in that genetic correlations were not perfect between ages 14, 17, and 24 in almost every case. This was true even for traits related to PEM, which did not significantly change phenotypically (on average) across assessments. Compared to genetic stability, environmental correlations were much smaller, meaning there was less consistency over time in the effects of individual life experiences on personality. However, these correlations were significantly higher than zero, suggesting some stability in these factors, and they increased with age to an average of .55 between ages 24 and 29.

Because the influences on personality change over time, and because these changes are not solely environmental, it is possible that including both young adults and older adults in the same genetic association studies may increase the difficulty in uncovering such associations. One way of exploring this issue is to divide samples by age and consider them separately, as I did in the current study. This does have the unfortunate effect of decreasing the sample size and therefore decreasing power to detect the small effects I expect. However, since GCTA considers the effects of all SNPs at once rather than individually, the multiple testing burden of GWAS is lifted and power increased. In the offspring sample, estimates of SNP heritability varied across assessments with large standard errors due to the limited sample size. Averaged across traits, estimates decreased from .18 at age 17 to .13 at age 29, suggesting that SNP heritability may decrease over time. In the adult sample, which was much larger, I observed significant estimates of SNP heritability in the traits aggression, harm avoidance, traditionalism, and the super-factor constraint. Averaged across those 4 traits, SNP heritability was .22. The average across all traits in adulthood was only .11 in contrast.

Another way of approaching the issue of developmental noise is to consider only what phenotypic variance is consistent over time. To provide an indication of that, I averaged trait scores across assessments. I observed that these composite phenotypes resulted in larger biometric univariate estimates of additive genetic effects, which suggests that GCTA may be more successful. SNP heritabilities of the composite phenotypes were still in many cases zero; in fact, none were statistically significant. However, I did observe an average across scales that was equal to that of the adults (M =

.11), and I observed estimates similar to those of the adults on the harm avoidance and traditionalism scales (explaining 20% and 21% of the variance of the composite phenotypes and 20% and 23% of the variance of the adult phenotypes, respectively) as well as the super-factor constraint (explaining 15% and 23% of the variance in the composite and adult phenotype, respectively). Even in adults, however, the estimates of SNP heritability were much smaller than estimates of additive effects from twin studies would suggest. In fact, common SNPs did not explain even half of the twin-estimated heritability.

Past research on adults or mixed samples have uncovered modest estimates of SNP heritability; regardless of measurement, these studies have produced estimates that are substantially lower than twin-estimated heritability for Big Five traits and their equivalents (de Moor et al., 2015; Power & Pluess, 2015; Vinkhuyzen et al., 2012), subjective well being (Rietveld et al., 2013), and TPQ harm avoidance, novelty seeking, reward dependence, and persistence (Verweij et al., 2012). On the other hand, studies addressing psychopathology have been somewhat more successful (MDD = .32, Lubke et al., 2012; behavioral disinhibition phenotypes = .08 - .37, McGue et al., 2013). If we consider only the results from the analyses of the parent generation, we see that it is not all that different from past research. Estimates for the super-factors NEM and PEM were not significantly different from zero, nor were those for most of the primary scales most highly related to them. However, the significant estimates observed were those for the constraint super-factor and its related primary scales harm avoidance and traditionalism, and for aggression. Similarly, the non-substance-related component of behavioral

disinhibition in McGue et al. (2013) produced an estimate of SNP heritability = .28; their sample overlaps greatly with the current sample, but I confirmed that separating adults and adolescents produces similar results. Behavioral disinhibition includes a cluster of personality traits such as low traditionalism and higher aggression (Benning et al., 2003). It may be the case that much of the genetic variance in extraversion- and neuroticism-related traits cannot be explained by common SNPs whereas a significant portion of the genetic variance in conscientiousness-related traits (or externalizing behavior) can be explained by common SNPs.

Although the results from the offspring generation are inconsistent and require replication, it appears that we can explain at least some of the heritability of personality traits related to behavioral disinhibition but that effects of common SNPs are in fact very small for other traits regardless of developmental stage of life. Still, it remains a possibility that development affects outcomes of GCTA at earlier ages. For example, Trzaskowski, Dale, and Plomin (2013) conducted GCTA on behavior problems in children aged 11-12; they found that they could not account for any of the heritability implied by their twin analyses. They suggested that they did not explain the twin heritability because much of it is non-additive and because they did not have enough power to detect smaller effects; I agree that those are both plausible explanations, however, I propose that an additional factor could be development, especially given the lower genetic correlations observed in childhood (Briley & Tucker-Drob, 2014).

I acknowledge the limitation of sample size in the GCTA of our developmental sample. It is possible that a lack of statistical power was responsible for the instability

and small size of individual h^2_{SNP} estimates. Power analysis (based on GREML power tool from GCTA website) suggests that for 80% power to detect SNP heritability for a trait with a true heritability of .6 (an upper-bound estimate for personality), sample size must be at least 1500, and even greater for a true heritability of .5 (~1800) or .4 (2000+). Although the current sample is large, it contains many related individuals, so the sample is essentially cut in half by restricting the relationship threshold to < .025. Age 14 was not feasible for GCTA as the sample size would have been about 600, which is not large enough to even detect very large effects. The age 17 assessment had the largest sample of offspring of any individual assessment (close to 1400), and this was marginally increased to ~1460 by considering COMP phenotypes. Thus, I may have had power to detect very large effects of common SNPs in the developmental samples but not enough to detect smaller effects.

Another limitation I faced is that, like numerous other personality studies, I relied on self-reports. This may particularly be an issue for younger participants, as the psychometric properties of such reports are less established in childhood and adolescence (Soto, John, Gosling, & Potter, 2008). However, I observed substantial estimates of internal consistency (Cronbach's alpha) for all primary scales at all ages; in fact, estimates at age 14 (.84 - .91) were no smaller than those at older ages (.78 - .92 across ages 17, 24, and 29). Still, the extent to which self-reports reflect actual behavior is not known and may differ over time.

What we have learned is that restricting GCTA to different developmental groups does not uncover a vast difference in SNP heritability over time; that is, I did not observe

larger h^2_{SNP} estimates than those observed in combined samples. However, I have not excluded the possibility of smaller developmental effects or of differing genetic effects in childhood or old age. If we continue to observe estimates of the influence of common SNPs that are well below those of total additive effects estimated from twin studies, we might conclude that the variants relevant to variation in personality are rare, not common, SNPs and/or there is substantial non-additivity (which is not well estimated in biometric studies of twins alone, but is suggested by the relatively smaller heritability estimates from parent-offspring studies; e.g., Bratko, Butkovic, Vukasovi, Kerestes, & Brkovic, 2012; Finkel & McGue, 1997; Vukasovic & Bratko, 2015). However, this does not necessarily apply to all traits. In the current analysis, a significant amount of the twin-estimated heritability was accounted for by SNP heritability for a few traits related to behavioral disinhibition. Future research should further explore SNP heritability of these traits.

As outlined in Verweij et al. (2013), there are different possible evolutionary explanations for the persistence of individual differences in personality. The 3 most popular theories entail differing underlying genetic architecture. First, selective neutrality (in which relevant genetic variants either have no effects on fitness or have an equal cost-benefit fitness ratio) implies high genetic additivity and low genetic non-additivity. Second, mutation-selection balance (in which deleterious alleles are constantly being selected out of the human genome while new ones are popping up) implies that relevant alleles should be rare and that there should be substantial non-additive effects on personality. Third, balancing selection (in which the fitness effects of particular alleles

differ depending on environmental context) implies high genetic additivity that is based on a relatively smaller number of important variants. Results of GWAS suggest that balancing selection is unlikely to be working for personality traits, but leaves open the possibility for either selective neutrality or mutation-selection balance. GCTA results suggest that for many traits, mutation-selection balance is the likelier mechanism of variation maintenance, as estimates of SNP heritability are much lower than biometric analyses would suggest. However, that may be the case for only some personality traits. In the current analysis, much of the twin-estimated heritability was accounted for by SNP heritability for traits related to behavioral disinhibition (constraint—harm avoidance and traditionalism, and aggression). It may be the case that, in terms of fitness, there are both costs and benefits of these traits. For example, being willing to take risks might increase an individual's chances of producing offspring, but at the same time may also decrease that individual's chances of survival. This would align with the selective neutrality hypothesis. Positive and negative emotionality, on the other hand, may be under a different kind of selection pressure—simply to weed out alleles that interfere with normal emotional processing.

Table 1. Description of primary scales of the Multidimensional Personality Questionnaire (MPQ).

Primary scale	Description of high scorers	Higher-order factor representation
Well-Being	Happy, cheerful, active, optimistic, positive self-concept	Positive Emotionality
Social Potency	Forceful, decisive, persuasive, enjoys leadership roles	Positive Emotionality
Achievement	Works hard, persists, likes long hours and demanding projects	Positive Emotionality
Social Closeness	Sociable, likes people, warm, affectionate	Positive Emotionality
Stress Reaction	Nervous, easily upset, troubled by guilt, feels vulnerable	Negative Emotionality
Alienation	Victim of bad luck, feels mistreated and betrayed	Negative Emotionality
Aggression	Physically aggressive, vindictive, likes violent scenes	Negative Emotionality
Control	Reflective, cautious, careful, rational, sensible	Constraint
Harm Avoidance	Does not enjoy excitement of adventure and danger	Constraint
Traditionalism	Endorses high moral standards, supports religious values	Constraint
Absorption	Emotionally responsive to engaging lights and sounds	Mixed

Table 2. Number of genotyped, Caucasian individuals and complete twin pairs with phenotypes at each target age.

		MZs	DZs	Total	Mean Age (SD)
14	**Individuals**	843	515	1358	14.91 (.55)
	Pairs	419	245	664	
17	**Individuals**	1858	1004	2862	17.85 (.51)
	Pairs	922	475	1397	
24	**Individuals**	1565	797	2362	24.98 (.86)
	Pairs	728	379	1107	
29	**Individuals**	1420	757	2177	29.48 (.58)
	Pairs	706	360	1066	
Adults	**Individuals**	-	-	3799	43.73 (5.61)

Table 3. Mean (SD) MPQ scale scores and N by age and sex.

	14		**17**		**24**		**29**		**ANOVA**		
	M	F	M	F	M	F	M	F			η^2
AB	-	-	42.09	43.18	40.85	40.05	39.24	38.48	Time	.107**	
			(8.90)	(9.65)	(9.12)	(9.26)	(9.47)	(9.31)	Sex	.000	
			1266	1504	954	1167	887	1167	T x S	.009*	
AC	-	-	49.61	48.10	52.90	50.06	53.21	50.52	Time	.073**	
			(7.91)	(8.48)	(7.60)	(7.75)	(7.90)	(8.02)	Sex	.029**	
			1264	1505	954	1167	887	1191	T x S	.005**	
AG	43.30	35.32	41.75	34.25	35.67	29.29	33.73	27.97	Time	.351**	
	(9.10)	(9.20)	(9.07)	(8.94)	(8.05)	(6.70)	(7.53)	(6.36)	Sex	.202**	
	376	958	1325	1507	898	1170	826	1190	T x S	.010**	
AL	37.76	35.44	35.71	34.59	31.31	30.10	30.30	29.15	Time	.231**	
	(8.55)	(9.27)	(8.41)	(9.06)	(8.17)	(8.00)	(8.24)	(8.23)	Sex	.008**	
	374	958	1327	1507	898	1170	827	1191	T x S	.000	
CN	45.15	47.29	46.60	47.93	49.38	52.26	50.99	53.47	Time	.207**	
	(7.84)	(8.05)	(7.14)	(7.95)	(7.30)	(7.44)	(7.43)	(7.61)	Sex	.031**	
	375	955	1326	1504	896	1171	826	1192	T x S	.005**	
HA	45.57	50.05	42.65	49.69	44.08	54.00	46.53	56.15	Time	.161**	
	(9.87)	(9.98)	(9.49)	(9.97)	(10.38)	(9.61)	(10.29)	(9.41)	Sex	.184**	
	369	950	1267	1509	957	1171	887	1189	T x S	.017**	
SC	-	-	52.60	55.91	52.52	56.60	52.24	56.45	Time	.002*	
			(7.53)	(8.66)	(8.08)	(8.10)	(8.16)	(8.35)	Sex	.061**	
			1270	1510	959	1172	888	1191	T x S	.004*	
SP	-	-	45.90	45.09	47.08	43.63	46.74	43.12	Time	.010**	
			(7.67)	(8.49)	(8.33)	(8.55)	(8.54)	(8.73)	Sex	.035**	
			1268	1505	957	1153	887	1191	T x S	.017**	
SR	39.94	42.71	41.26	44.57	38.61	42.05	37.91	41.73	Time	.095**	
	(8.15)	(9.53)	(8.61)	(9.54)	(8.54)	(9.24)	(8.88)	(9.36)	Sex	.040**	
	374	960	1266	1506	955	1169	887	1190	T x S	.002*	
TR	-	-	50.69	51.91	52.05	52.34	52.87	52.51	Time	.008**	
			(6.70)	(6.99)	(7.45)	(6.82)	(6.80)	(6.92)	Sex	.000	
			1256	1456	947	1140	882	1182	T x S	.008**	
WB	54.96	56.27	54.89	55.73	55.56	55.94	55.37	56.15	Time	.002	
	(7.33)	(8.57)	(7.23)	(8.39)	(7.36)	(8.12)	(7.48)	(8.11)	Sex	.000	
	375	960	1264	1504	955	1168	885	1190	T x S	.000	
PEM	-	-	123.43	122.72	125.41	122.10	124.59	121.36	Time	.005**	
			(12.10)	(14.25)	(12.82)	(13.28)	(13.22)	(13.23)	Sex	.012**	
			1250	1483	940	1151	879	1176	T x S	.007**	
NEM	-	-	90.01	87.70	83.28	80.17	81.32	78.11	Time	.266**	
			(13.67)	(14.67)	(13.41)	(13.48)	(13.70)	(13.36)	Sex	.016**	
			1250	1483	940	1162	879	1165	T x S	.000	
CON	-	-	129.43	137.48	135.02	145.21	139.02	148.09	Time	.217**	
			(14.90)	(15.86)	(15.36)	(14.30)	(14.80)	(14.65)	Sex	.096**	
			1250	1483	940	1151	879	1176	T x S	.002*	

Note. Some scales are missing data because they are not included in the PBYA given to 14-year-olds.

*p < .05, ** p < .001

Table 4. Rank order stability (Pearson r and CIs) of all traits across assessments.

	14-17	14-24	14-29	17-24	17-29	24-29
AB	-	-	-	.56 (.53-.59)	.52 (.49-.55)	.74 (.72-.76)
AC	-	-	-	.53 (.49-.56)	.50 (.47-.53)	.70 (.67-.72)
AG	.62 (.58-.66)	.49 (.42-.55)	.46 (.39-.52)	.60 (.57-.63)	.58 (.55-.61)	.74 (.72-.76)
AL	.54 (.50-.58)	.47 (.40-.53)	.41 (.34-.48)	.51 (.48-.54)	.49 (.45-.52)	.70 (.68-.73)
CN	.56 (.52-.60)	.37 (.29-.44)	.39 (.32-.46)	.52 (.49-.55)	.48 (.45-.52)	.73 (.71-.76)
HA	.62 (.58-.66)	.46 (.39-.53)	.48 (.41-.54)	.68 (.66-.71)	.65 (.62-.68)	.82 (.81-.84)
SC	-	-	-	.54 (.51-.57)	.52 (.49-.56)	.75 (.73-.77)
SP	-	-	-	.61 (.59-.64)	.58 (.55-.61)	.78 (.77-.80)
SR	.60 (.56-.64)	.51 (.45-.57)	.48 (.41-.54)	.56 (.53-.59)	.53 (.50-.56)	.72 (.70-.75)
TR	-	-	-	.53 (.50-.56)	.50 (.47-.54)	.74 (.71-.76)
WB	.55 (.51-.59)	.43 (.36-.50)	.41 (.34-.48)	.48 (.45-.52)	.47 (.44-.51)	.70 (.67-.72)
PEM	-	-	-	.57 (.54-.60)	.56 (.53-.59)	.74 (.72-.76)
NEM	-	-	-	.55 (.52-.58)	.53 (.49-.56)	.74 (.72-.76)
CON	-	-	-	.62 (.59-.64)	.58 (.55-.61)	.79 (.78-.81)

Note. Retest correlations cannot be calculated between age 14 and 17, 24, or 29 for some scales because those scales are not included in the PBYA given to 14-year-olds.

Table 5. Twin correlations and standardized parameter estimates (CIs) from the univariate Cholesky decomposition model.

		rMZ	rDZ	a^2	c^2	e^2
AB	**17**	**.50 (.45-.55)**	**.21 (.13-.29)**	**.49 (.37-.54)**	.00 (0-.11)	**.51 (.46-.55)**
	24	**.46 (.40-.52)**	**.24 (.14-.33)**	**.44 (.22-.51)**	.02 (0-.21)	**.54 (.49-.60)**
	29	**.45 (.39-.51)**	**.23 (.13-.32)**	**.43 (.22-.50)**	.02 (0-.21)	**.55 (.50-.61)**
	COMP	**.55 (.51-.60)**	**.23 (.15-.31)**	**.56 (.36-.60)**	.00 (0-.18)	**.44 (.40-.50)**
AC	**17**	**.45 (.40-.50)**	**.14 (.05-.22)**	**.43 (.35-.48)**	.00 (0-.07)	**.57 (.52-.62)**
	24	**.45 (.39-.51)**	**.14 (.03-.23)**	**.45 (.36-.51)**	.00 (0-.07)	**.55 (.49-.61)**
	29	**.44 (.38-.50)**	**.14 (.04-.23)**	**.44 (.34-.49)**	.00 (0-.08)	**.56 (.51-.63)**
	COMP	**.51 (.47-.56)**	**.13 (.05-.21)**	**.51 (.43-.56)**	.00 (0-.06)	**.49 (.44-.55)**
AG	**14**	**.52 (.46-.59)**	**.38 (.28-.47)**	**.31 (.10-.53)**	**.22 (.02-.40)**	**.47 (.41-.54)**
	17	**.59 (.55-.63)**	**.33 (.25-.40)**	**.50 (.34-.62)**	.09 (0-.23)	**.42 (.38-.46)**
	24	**.58 (.53-.63)**	**.36 (.27-.45)**	**.44 (.25-.62)**	.15 (0-.31)	**.42 (.37-.47)**
	29	**.53 (.47-.58)**	**.34 (.24-.43)**	**.39 (.19-.57)**	.15 (0-.32)	**.47 (.42-.52)**
	COMP	**.67 (.64-.71)**	**.40 (.33- .47)**	**.63 (.48-.71)**	.06 (0-.20)	**.32 (.28-.35)**
AL	**14**	**.57 (.50-.63)**	**.32 (.21-.42)**	**.51 (.30-.63)**	.04 (0-.23)	**.43 (.38-.50)**
	17	**.50 (.45-.55)**	**.29 (.21-.37)**	**.41 (.24-.54)**	.00 (0-.24)	**.49 (.45-.55)**
	24	**.45 (.39-.51)**	**.28 (.18-.37)**	**.37 (.16-.51)**	.08 (0-.26)	**.55 (.49-.61)**
	29	**.46 (.39-.52)**	**.25 (.15-.35)**	**.42 (.21-.52)**	.03 (0-.33)	**.54 (.45-.63)**
	COMP	**.60 (.56-.64)**	**.39 (.31-.46)**	**.45 (.29-.62)**	.16 (0-.31)	**.40 (.36-.44)**
CN	**14**	**.40 (.32-.48)**	**.20 (.09-.31)**	**.42 (.22-.49)**	.00 (0-.17)	**.58 (.51-.66)**
	17	**.39 (.33-.44)**	.05 (-.03-.13)	**.37 (.31-.42)**	.00 (0-.04)	**.63 (.58-.69)**
	24	**.34 (.27-.41)**	.10 (0-.21)	**.35 (.25-.41)**	.00 (0-.07)	**.65 (.59-.72)**
	29	**.32 (.25-.38)**	.08 (-.02-.18)	**.31 (.20-.38)**	.00 (0-.08)	**.69 (.62-.76)**
	COMP	**.47 (.41-.51)**	**.13 (.05-.21)**	**.44 (.37-.49)**	.00 (0-.05)	**.56 (.51-.62)**
HA	**14**	**.60 (.54-.66)**	**.17 (.05-.28)**	**.59 (.50-.64)**	.00 (0-.07)	**.41 (.36-.47)**
	17	**.58 (.54-.62)**	**.29 (.21-.36)**	**.58 (.42-.62)**	.00 (0-.15)	**.42 (.38-.46)**
	24	**.65 (.60-.69)**	**.42 (.34-.50)**	**.50 (.33-.67)**	.16 (0-.31)	**.34 (.30-.39)**
	29	**.62 (.57-.67)**	**.37 (.28-.45)**	**.57 (.39-.68)**	.07 (0-.23)	**.36 (.32-.41)**
	COMP	**.71 (.68-.74)**	**.39 (.32-.46)**	**.66 (.52-.74)**	.05 (0-.18)	**.29 (.26-.32)**
SC	**17**	**.45 (.39-.49)**	**.12 (.04-.20)**	**.44 (.37-.49)**	.00 (0-.05)	**.56 (.51-.61)**
	24	**.49 (.43-.54)**	**.20 (.10-.29)**	**.49 (.37-.54)**	.00 (0-.10)	**.51 (.46-.57)**
	29	**.49 (.43-.55)**	**.17 (.07-.26)**	**.49 (.39-.54)**	.00 (0-.08)	**.51 (.46-.57)**
	COMP	**.55 (.50-.59)**	**.20 (.12-.38)**	**.54 (.44-.59)**	.00 (0-.09)	**.46 (.41-.51)**
SP	**17**	**.49 (.44-.54)**	**.13 (.04-.21)**	**.47 (.41-.52)**	.00 (0-.05)	**.53 (.48-.57)**
	24	**.46 (.40-.52)**	**.23 (.13-.32)**	**.47 (.30-.52)**	.00 (0-.14)	**.53 (.48-.59)**
	29	**.50 (.44-.56)**	**.21 (.11-.31)**	**.50 (.37-.55)**	.00 (0-.11)	**.50 (.45-.56)**
	COMP	**.55 (.50-.59)**	**.21 (.12-.28)**	**.58 (.46-.62)**	.00 (0-.10)	**.42 (.38-.47)**
SR	**14**	**.50 (.43-.57)**	**.23 (.12-.34)**	**.50 (.29-.56)**	.00 (0-.18)	**.50 (.44-.57)**
	17	**.45 (.39-.50)**	**.22 (.13-.30)**	**.44 (.27-.49)**	.00 (0-.15)	**.56 (.51-.61)**
	24	**.44 (.37-.49)**	**.21 (.12-.31)**	**.44 (.24-.49)**	.00 (0-.17)	**.56 (.51-.62)**
	29	**.42 (.35-.48)**	**.22 (.12-.32)**	**.41 (.19-.48)**	.02 (0-.20)	**.57 (.52-.64)**
	COMP	**.54 (.49-.58)**	**.27 (.19-.34)**	**.54 (.38-.59)**	.00 (0-.15)	**.46 (.41-.50)**
TR	**17**	**.57 (.53-.61)**	**.36 (.29-.44)**	**.43 (.28-.60)**	.14 (0-.28)	**.42 (.39-.47)**
	24	**.59 (.54-.64)**	**.39 (.30-.47)**	**.45 (.28-.63)**	.15 (0-.30)	**.40 (.36-.45)**
	29	**.56 (.51-.61)**	**.38 (.29-.46)**	**.45 (.27-.62)**	.13 (0-.29)	**.42 (.37-.47)**

COMP	**.66 (.62-.69)**	**.43 (.35-.49)**	**.58 (.41-.70)**	.09 (0-.25)	**.34 (.30-.38)**
WB 14	**.49 (.42-.55)**	**.25 (.14-.35)**	**.49 (.25-.55)**	.00 (0-.21)	**.51 (.45-.58)**
17	**.50 (.45-.55)**	**.20 (.11-.28)**	**.50 (.38-.54)**	.00 (0-.10)	**.50 (.46-.55)**
24	**.42 (.36-.48)**	**.22 (.12-.31)**	**.42 (.20-.48)**	.01 (0-.20)	**.57 (.52-.64)**
29	**.44 (.37-.50)**	**.23 (.13-.32)**	**.40 (.18-.49)**	.04 (0-.23)	**.56 (.51-.63)**
COMP	**.62 (.58-.66)**	**.38 (.31-.45)**	**.49 (.32-.60)**	.07 (0-.23)	**.44 (.40-.48)**
PEM 17	**.53 (.48-.58)**	**.22 (.14-.30)**	**.52 (.39-.57)**	.00 (0-.12)	**.48 (.43-.52)**
24	**.45 (.38-.51)**	**.23 (.13-.33)**	**.47 (.29-.52)**	.00 (0-.15)	**.53 (.48-.59)**
29	**.50 (.46-.56)**	**.21 (.11-.31)**	**.50 (.36-.56)**	.00 (0-.12)	**.42 (.37-.47)**
COMP	**.59 (.55-.63)**	**.29 (.21-.37)**	**.58 (.46-.63)**	.00 (0-.11)	**.42 (.37-.47)**
NEM 17	**.49 (.44-.53)**	**.19 (.11-.27)**	**.48 (.34-.52)**	.00 (0-.12)	**.52 (.48-.57)**
24	**.48 (.41-.53)**	**.26 (.16-.35)**	**.44 (.23-.53)**	.04 (0-.22)	**.53 (.47-.59)**
29	**.45 (.38-.51)**	**.24 (.14-.33)**	**.46 (.25-.52)**	.00 (0-.18)	**.54 (.48-.60)**
COMP	**.55 (.50-.59)**	**.28 (.21-.36)**	**.56 (.36-.61)**	.00 (0-.18)	**.44 (.39-.49)**
CON 17	**.56 (.51-.60)**	**.23 (.15-.31)**	**.55 (.46-.59)**	.00 (0-.08)	**.45 (.41-.49)**
24	**.58 (.53-.63)**	**.27 (.17-.36)**	**.59 (.47-.63)**	.00 (0-.11)	**.41 (.37-.46)**
29	**.53 (.47-.58)**	**.22 (.12-.32)**	**.54 (.44-.59)**	.00 (0-.09)	**.46 (.41-.51)**
COMP	**.62 (.58-.66)**	**.28 (.20-.36)**	**.63 (.54-.67)**	.00 (0-.08)	**.37 (.33-.42)**

Note. Some scales are missing results because such scales are not included in the PBYA given to 14-year-olds. COMP = composite phenotypes. Bold text indicates statistical significance.

Table 6. Genetic and environmental correlations (CIs) from the bivariate Cholesky decomposition model.

		14-17	**14-24**	**14-29**	**17-24**	**17-29**	**24-29**
AB	r_g r_e	-	-	-	.92 (.92-1.0) .34 (.27-.40)	.84 (.68-1.0) .30 (.23-.37)	.95 (.88-1.0) .57 (.52-.62)
AC	r_g r_e	-	-	-	.73 (.65-.99) .38 (.32-.44)	.73 (.62-1.0) .34 (.28-.41)	1.00 (.92-1.0) .48 (.42-.53)
AG	r_g r_e	1.00 (.84-1.0) .39 (.31-.47)	1.00 (.73-1.0) .22 (.08-.35)	.92 (.42-1.0) .34 (.20-.48)	.79 (.68-1.0) .31 (.24-.38)	.84 (.64 - 1.0) .32 (.25-.39)	.95 (.87-1.0) .51 (.45-.57)
AL	r_g r_e	.67 (.47-.86) .35 (.26-.43)	.69 (.46-1.0) .24 (.09-.37)	.93 (.50-1.0) .22 (.09-.35)	.69 (.50-.93) .34 (.28-.41)	.73 (.54-1.0) .27 (.19-.33)	.93 (.84-1.0) .52 (.46-.57)
CN	r_g r_e	.89 (.67-1.0) .41 (.34-.48)	.63 (.40-.92) .27 (.14-.39)	.67 (.45-1.0) .29 (.16-.41)	.89 (.73-1.0) .40 (.33-.46)	.71 (.60-.87) .39 (.32-.45)	.91 (.84-1.0) .64 (.60-.68)
HA	r_g r_e	.84 (.77-.99) .34 (.26-.42)	.76 (.56-1.0) .27 (.12-.40)	.77 (.61-1.0) .29 (.15-.42)	.96 (.86-1.0) .34 (.27-.40)	.86 (.79-.98) .34 (.27-.40)	1.00 (.95-1.0) .54 (.49-.59)
SC	r_g r_e	-	-	-	.78 (.70-.86) .36 (.30-.42)	.76 (.68-.83) .33(.26-.39)	.97 (.89-1.0) .57 (.51-.62)
SP	r_g r_e	-	-	-	.82 (.75-.99) .45 (.39-.50)	.79 (.70-.99) .41 (.35-.47)	.99 (.95-1.0) .59 (.55-.64)
SR	r_g r_e	.82 (.70-1.0) .39 (.31-.46)	.83 (.53-1.0) .32 (.19-.43)	.78 (.56-1.0) .24 (.11-.36)	.84 (.69-1.0) .37 (.31-.44)	.83 (.63-1.0) .35 (.28-.41)	.97 (.91-1.0) .56 (.51-.61)
TR	r_g r_e	-	-	-	.77 (.58-1.0) .25 (.18-.32)	.65 (.49-.90) .23 (.15-.30)	1.00 (.89-1.0) .43 (.37-.49)
WB	r_g r_e	.75 (.60-.95) .33 (.24-.41)	.70 (.44-1.0) .25 (.10-.37)	.45 (.04-.81) .28 (.15-.40)	.76 (.62-.93) .26 (.19-.32)	.75 (.62-.98) .23 (.16-.29)	1.00 (.92-1.0) .51 (.46-.56)
PEM	r_g r_e	-	-	-	.78 (.68-.90) .37 (.31-.43)	.70 (.61-.81) .39 (.33-.45)	.96 (.91-1.0) .55 (.49-.60)
NEM	r_g r_e	-	-	-	.77 (.63-1.0) .37 (.31-.43)	.80 (.67-1.0) .32 (.25-.39)	.95 (.88-1.0) .59 (.53-.63)
CON	r_g r_e	-	-	-	.94 (.78-1.0) .49 (.43-.54)	.79 (.73-.89) .37 (.30 -.44)	.95 (.92-1.0) .60 (.55-.65)
Mean	r_g r_e	.83 .37	.78 .26	.75 .28	.81 .35	.77 .32	.97 .54

Note. Some scales are missing results because such scales are not included in the PBYA given to 14-year-olds. r_g = genetic correlation, r_e = environmental correlation.

Table 7. Univariate SNP heritability estimates (SE) and N from GCTA at different ages.

	17	**24**	**29**	**COMP**	**Adults**
AB	.00 (.26) 1390	.00 (.31) 1121	.04 (.33) 1083	.00 (.25) 1457	.08 (.10) 3383
AC	.05 (.25) 1389	.00 (.31) 1120	.00 (.32) 1084	.00 (.24) 1457	.00 (.10) 3384
AG	.00 (.25) 1388	.24 (.33) 1097	.08 (.33) 1061	.00 (.24) 1457	.22 (.11) 3385
AL	.41 (.26) 1389	.14 (.33) 1098	.01 (.34) 1061	.18 (.25) 1457	.04 (.10) 3385
CN	.07 (.25) 1388	.05 (.32) 1097	.00 (.34) 1062	.00 (.24) 1457	.12 (.10) 3387
HA	.17 (.26) 1389	.20 (.33) 1119	.12 (.32) 1084	.20 (.24) 1457	.20 (.10) 3383
SC	.19 (.26) 1389	.00 (.32) 1120	.00 (.31) 1084	.00 (.24) 1457	.07 (.10) 3388
SP	.27 (.26) 1389	.20 (.32) 1113	.47 (.33) 1084	.24 (.25) 1457	.13 (.11) 3386
SR	.32 (.25) 1389	.60 (.31) 1119	.12 (.32) 1083	.35 (.24) 1457	.10 (.11) 3385
TR	.47 (.25) 1387	.08 (.32) 1111	.32 (.32) 1083	.21 (.23) 1456	.23 (.10) 3364
WB	.03 (.25) 1387	.19 (.31) 1120	.23 (.32) 1082	.13 (.24) 1457	.01 (.10) 3383
PEM	.00 (.25) 1386	.27 (.32) 1116	.43 (.33) 1080	.00 (.24) 1456	.00 (.11) 3340
NEM	.27 (.26) 1386	.20 (.32) 1116	.05 (.33) 1075	.03 (.25) 1456	.08 (.11) 3340
CON	.25 (.25) 1386	.00 (.32) 1116	.00 (.32) 1080	.15 (.24) 1456	.23 (.10) 3340
Average	.18	.15	.13	.12	.11

Note. Sample size was not feasible for GCTA at age 14. Average = mean of primary scales only.

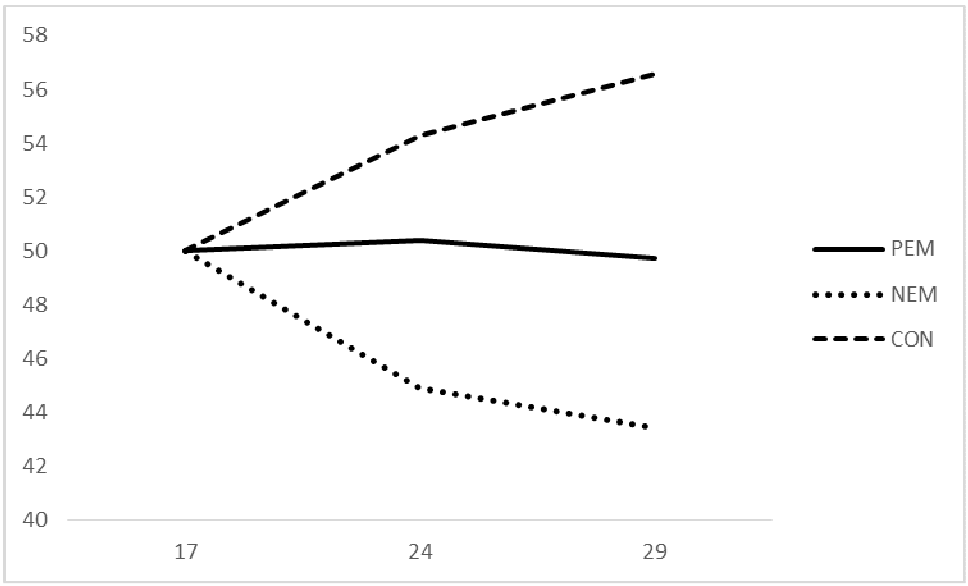

Figure 1. Change in T-scores over time for 3 super-factors.
Note. Age 14 is not included in this figure because the super-factors are not able to be calculated from the limited scales of the PBYA. Scores were scaled so that $M = 50$ and $SD = 10$ at age 17.

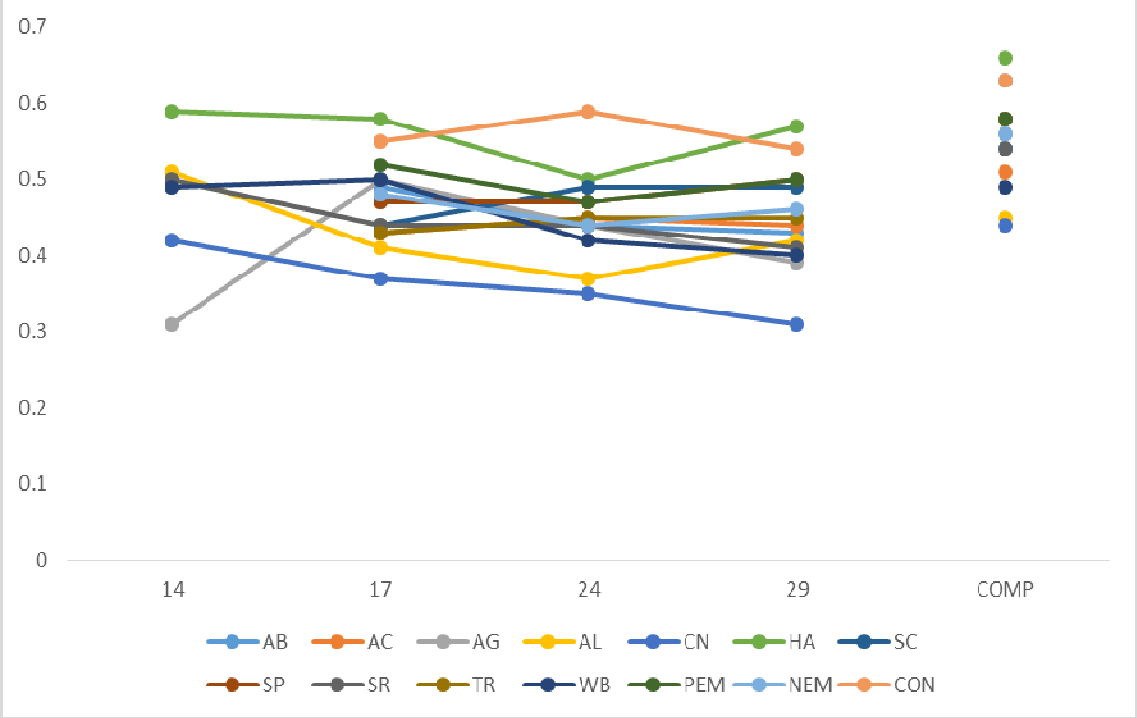

Figure 2. Additive effects (a^2) of all traits by age in the offspring sample, including the composite phenotype (COMP).
Note. Estimates are based on biometrical analysis of twin data.

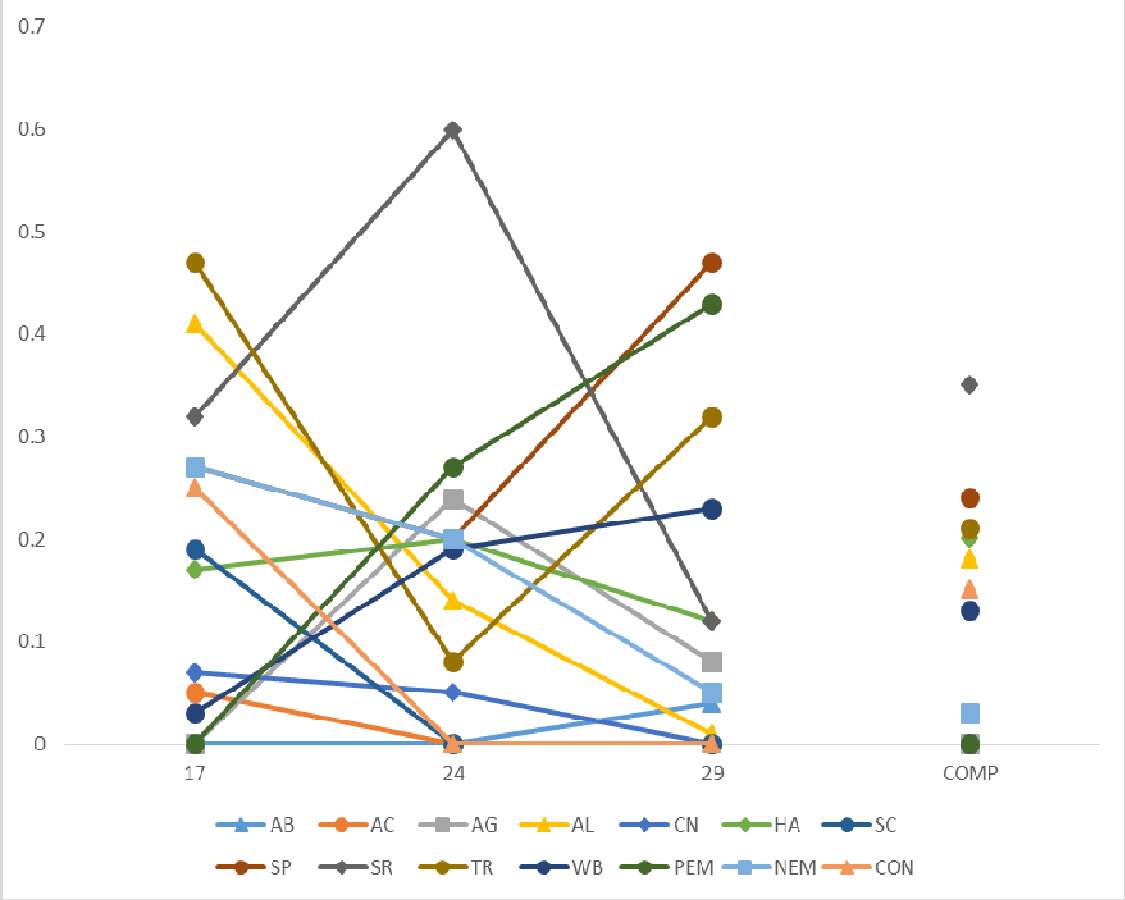

Figure 3. SNP heritability of all traits across age in the offspring sample, including the composite phenotype (COMP).
Note. Sample size was not feasible for GCTA at age 14.

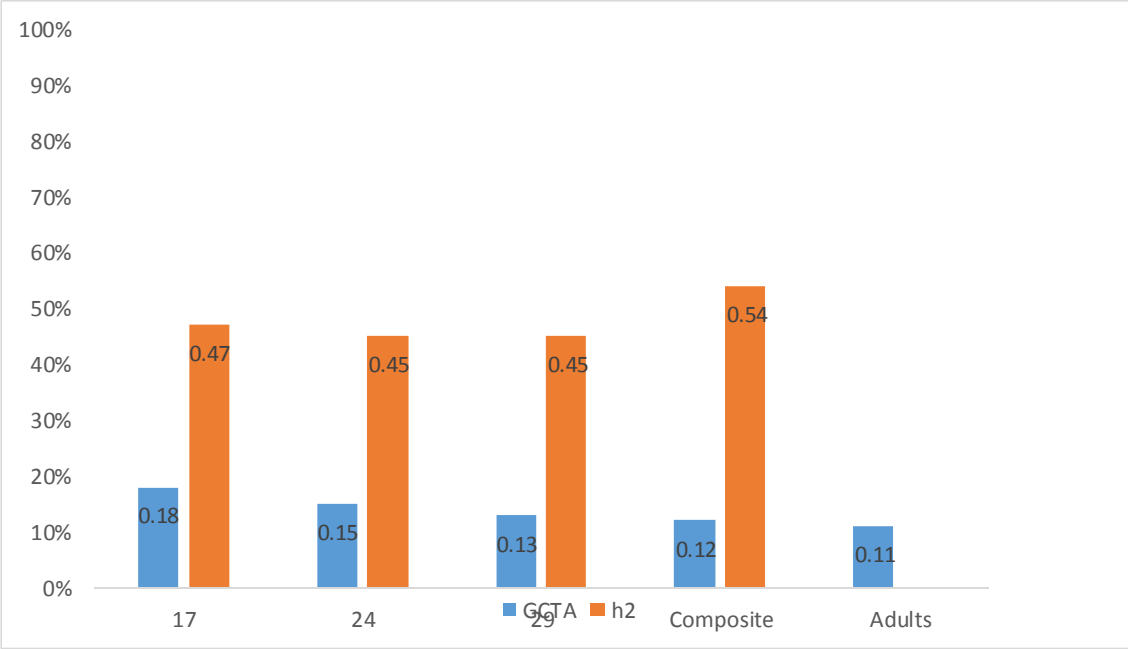

Figure 4. Average heritability estimates, GCTA vs. twin model.
Note. Sample size was not feasible for GCTA at age 14.

Chapter 3. "Re-hatching" the Chicken and Egg Problem: Does Biologically Influenced Personality Change Allow Young Adults to Take On New Social Roles, or Do New Social Roles Inspire Personality Change?

Decades of research findings have converged on some undeniable facts regarding the development of personality. The first is that personality is incredibly consistent; there is stability across measurement as well as across time. With respect to measurement, in a recent meta-analysis Connelly and Ones (2010) found that, though the correlations varied by trait, outside observers' ratings of individuals' personalities were significantly correlated with 1) other observers' ratings, 2) the individual's self-ratings, and 3) criterion measures (strangers' first impressions, academic achievement, and job performance). With respect to time, retest correlations are large (> .4) even across several years (e.g., Roberts & DelVecchio, 2000). These studies have provided evidence of the reliability and validity of trait constructs as persistent and pervasive characteristics of individuals. Related to this is a second important fact, namely, that personality stability increases as people age. In a large meta-analysis, Roberts and DelVecchio (2000) reported 7-year retest correlations of .43 in adolescence, .6 in young adulthood, and .74 in old age. This pattern of increasing retest correlations has been referred to as the *Cumulative Continuity Principle* (Caspi, Roberts, & Shiner, 2005).

Despite its substantial consistency, personality is by no means a static attribute of individuals, and change can be conceptualized in different ways. Research typically distinguishes between differences in rank order and differences in absolute scores over

time. The retest correlations previously mentioned represent the degree to which individuals keep their place relative to others in the population over a number of years and as such are referred to as stability coefficients. Although they are large, they do not reach unity at any age (Briley & Tucker-Drob, 2014). On the other hand, absolute change represents the degree to which scores actually change over time and can be examined individually or normatively (i.e., at the mean level). Another point of convergence in personality research concerns this latter type of change. The usual pattern of personality change can be summarized by the *Maturity Principle* (Caspi et al., 2005), which suggests that people change in ways that allow them to become functioning members of society. Specifically, research has shown that the greatest amount of personality change occurs during the transition from adolescence to adulthood and that people typically become more conscientious, agreeable, and socially dominant as well as less neurotic across this period of time (e.g., Roberts, Walton, & Viechtbauer, 2006). However, it is important to realize that there are substantial individual differences in the timing of change and even in its direction such that some people defy the norms.

What remains contentious is the explanation for these patterns of personality development. According to the Five Factor Theory (FFT) of personality, traits are internal, biologically influenced dispositions that drive behavior (McCrae et al., 2000). Proponents of this theory specify a developmental difference between basic tendencies (dispositions) and their behavioral and psychological manifestations (characteristic adaptations such as goals, skills, habits, etc.) such that environmental factors may affect only the expression of dispositions but not the dispositions themselves. This theory thus

implies that biological processes are responsible for stable individual differences in personality and that change is due to the effects of biologically-programmed events such as puberty (intrinsic maturation). It predicts that major personality changes will occur during some critical periods in life, like adolescence and old age, but that it will be highly stable in the meantime.

Social Investment Theory (SIT; Roberts, Wood, & Smith, 2005), on the other hand, suggests that it is investment in social roles that is responsible for personality development after adolescence. It proposes that personality changes will occur as a result of social role transitions (such as becoming a parent) and that personality will stabilize as the occurrence of major changes becomes less frequent and the environment increases in consistency. Specifically, each new social role presents a set of expectations, demands, challenges, rules, etc. and success in such roles requires aligning one's behavior to these expectations. Mean-level changes in personality thus reflect the tendency of most individuals to adapt to role changes in a similar way, for example, by becoming more responsible upon entering the workforce. SIT predicts that people will experience the greatest amount of personality change during young adulthood when most of these major role transitions occur, but as investment in these new roles increases, environment will stabilize and therefore so will personality.

One approach to testing these opposing theories is to use behavior genetic methodology; twin and family studies allow for the teasing apart of the relative influence of biological and social factors on personality development. Significant heritability has been observed at every age, and genetic factors consistently contribute to stability of

personality across the life span (Briley & Tucker-Drob, 2014). At the same time, genetic factors also contribute to personality change, at least in childhood, adolescence, and young adulthood. These facts support the intrinsic maturation tenet of FFT. However, these results do not necessarily disagree with SIT. The *Corresponsive Principle* of personality development (Caspi et al., 2005) states that the most likely effect of experiences is to reinforce through a feedback loop the characteristics that lead people to those experiences in the first place. This suggests a correlation between personality and environments, and to the extent that personality is heritable, a gene-environment correlation exists that inflates heritability estimates of personality. The findings that environmental factors do appear to significantly influence personality, that this influence increases across the life span, and that environment increasingly contributes to the stability of personality over time (Briley & Tucker-Drob, 2014) support SIT but are problematic for FFT. However, while results from biometric modeling inform developmental theory, a major limitation is that they do not identify the specific biological or environmental factors relevant to personality development. Further understanding of personality etiology requires examining the causal influence of biological and environmental factors on personality across the life span. In particular, a stronger argument for SIT could be made by demonstrating a causal effect of the occurrence (and timing) of major life transitions on personality change.

There is no shortage of research on the relationship between personality development and life events, the latter of which can refer to singular events (e.g., major illness or injury) or aggregates of events, short-term fluctuations in environmental

factors, or role transitions. While much of the previous research has focused on the occurrence of aggregate positive and negative life events (e.g., Jeronimus, Ormel, Aleman, Penninx, & Riese, 2013; Jeronimus, Riese, Sanderman, & Ormel, 2014; Kandler, Bleidorn, Riemann, Angleitner, & Spinath, 2012; Ludtke, Roberts, Trautwein, & Nagy, 2011; Specht, Egloff, & Schmukle, 2011), I propose that examination of single events may provide more information relevant to personality change because, as Jeronimus and colleagues (2013) noted, it is difficult to classify events as simply positive or negative for multiple reasons. First, some events may be considered both positive and negative. Second, positive and negative events may occur simultaneously and therefore their effects may interact or cancel out entirely. Third, the context in which an event occurs may have a large influence on whether the event is considered positive or negative. For example, an accidental pregnancy may not be experienced as positively as a planned pregnancy, and parenting is likely to be experienced as both positive and negative at times by everyone. In a subsequent study, Jeronimus and colleagues (2014) included more contextual information in the classification of the emotional valence of events; however, this information was derived largely from subjective recall of the experiences by the participants. An advantage of studying single events, and specific role transitions in particular, is that they do not need to be classified as positive or negative for the purpose of categorization into groups; also, they are likely to be linked to a specific date and thus not as influenced by the biases of retrospective reporting that may be involved when recalling the positive and negative experiences of life events.

Importantly, we may directly address SIT by examining the relationship between personality maturation and the experience of specific social roles; I have chosen life events that represent common milestones in young adulthood. Such events include first marriage and becoming a parent for the first time. Events are normative to the extent that they are expected to occur within a particular age as well as to the extent that they imply specific expectations for behavior (Neyer, Mund, Zimmermann, & Wrzus, 2014), which does seem to be the case with marriage and becoming a parent. Neyer et al. suggest that because normative experiences occur for a majority of individuals, they are expected to be less affected by self-selection; also, because normative experiences are supplied with behavioral guidelines, they are expected to have a larger influence on personality change than less normative experiences, which are thought to contribute more to the stability of personality through selection-socialization correlations (Corresponsive Principle). Thus, normative events should exhibit the strongest causal effects we might observe for life events.

While many studies suggest that life events are important to the development of personality, they are limited in that they are not able to determine causality. The same studies examining the impact of events on personality find that personality significantly predicts the events themselves. For example, neuroticism has been shown repeatedly to predict the occurrence of negative life events while extraversion has been shown to predict positive life events, and openness has been shown to predict both negative and positive life events (e.g., Kandler et al., 2012; Ludtke et al., 2011; Specht et al., 2011). Traits have also been shown to predict the occurrence of normative life events. In one

study, those who were less neurotic, more extraverted, and more conscientious were more likely to have started their first committed relationship (Neyer & Lehnart, 2007). In another study, those who were more satisfied with their lives were more likely to be married and to have children (Luhmann, Lucas, Eid, & Diener, 2013). In yet another study, those who were more conscientious were also more likely to be currently working and in a relationship (Leikas & Salmela-Aro, 2015). These studies provide evidence for selection effects that need to be taken into account in the consideration of the impact of life events. Though researchers have taken different approaches to accounting for selection effects, they still are not able to control for all possible alternative explanations for personality-event relationships.

One particular cause for concern is the possible confound of shared familial etiology (genetic and environmental) between personality and events. Past research has shown that life events are stable and more importantly heritable, not only in biometric (e.g., Kendler & Baker, 2007) but also in molecular (Power et al., 2013) studies. Furthermore, such heritability substantially overlaps with that of personality (Billig, Hershberger, Iacono, & McGue, 1996; Saudino, Pedersen, Lichtenstein, McClearn, & Plomin, 1997). I approach the issue of confounding by using the co-twin control method. In the co-twin control research design, it is possible to test whether the observation of a relationship holds up after controlling for factors shared by siblings. The logic of this method centers on the notion that because they are genetically identical and raised in the same home, monozygotic (MZ) twins provide the closest possible approximation to observing outcomes (in this case, personality change) in alternative exposure conditions

(i.e., event occurs or does not occur). That is, comparison of the personalities of identical twins who become discordant for incidence of an event is analogous to comparison of a single individual's personality trajectory after having both experienced and not experienced the event. Though it cannot completely isolate the event as the only possible causal factor in the analyses, comparing twins discordant for events can exclude the possibility that factors shared between MZs (including genes and environment) explain the effects of the events on personality change (see McGue, Osler, & Christensen, 2010 for a more thorough description of the counterfactual model for evaluating causal inference in epidemiology). Previous studies of this kind have demonstrated causal effects of marriage on desistence from antisocial behavior in men (Burt, Donnellan, Humbad, Hicks, McGue, & Iacono, 2010) and stressful life events and trauma on increased neuroticism and depression (Brown et al., 2014; Middeldorp Cath, Beem, Willemsen, & Boomsma, 2008).

In the current study, I examined personality with linear mixed models (with family and individual as random effects) as a function of age (i.e., 17 or 29), sex, and life event occurrence by age 29. I was particularly interested in the interaction of age and life event group, as significant values would suggest an effect of life events on personality change. Furthermore, I was interested in determining whether any observed effects of life events could be explained by familial factors; thus, I performed the co-twin control analysis with mixed models that included age and sex as well as zygosity and pair-level occurrence of life events (i.e., twins were assigned a group based on their marriage or parent status as well as that of their twin). In this analysis, the comparison of discordant

twins was of primary interest; significant differences between such groups would be consistent with a causal effect of life events. Additionally, the 3-way interaction of age, life event group, and zygosity was of interest, as significant values would suggest differential effects of life events on personality change depending on zygosity. Specifically, I aimed to answer the following questions:

1. Within our sample, do individuals' personalities grow in the direction of greater maturity (increased conscientiousness, agreeableness, and social dominance and decreased neuroticism) from age 17 to 29 and what is the magnitude of this change?
2. To what extent does the magnitude and direction of personality change from age 17 to 29 depend on the occurrence of major life transitions?
3. Is there evidence that the relationship between personality maturation and the occurrence of major life transitions is causal in nature?

Method

Participants

The current study included a sample of twins from the Minnesota Twin and Family Study (MTFS; Iacono & McGue, 2002). The MTFS is an ongoing longitudinal study of the etiological bases of substance abuse and related psychopathology. It includes both MZ and dizygotic (DZ) same-sex twins raised in the same home. Twins were recruited from publicly available Minnesota birth certificates; exclusion criteria included living further than a day's drive from Minneapolis as well as the presence of any mental

or physical handicap that would prevent an individual from completing the assessment. The sample is mostly Caucasian (over 95%) which is generally representative of Minnesota's demographics for the cohorts sampled. Zygosity of twins was assessed in three ways; determination was made on the basis of staff opinion of the twins' physical similarity, parent report on the Physical Similarity Questionnaire (Peeters, van Gestel, Vlietinck, Derom, & Derom, 1998), and an algorithm of physical measurement. Disagreements were resolved with a serological blood analysis.

Twins began participation at approximately age 11 or 17 and have returned on average every three or four years. Overall, about 90% of the sample has returned at each of the follow-ups. In the current analysis, personality data were obtained from assessments targeted at ages 17 and 29 whereas life events data were obtained from assessments targeted at ages 20, 24, and 29. Age 17 personality was available for 2456 individuals. I removed 24 individuals who reported being married or having children at age 17, 415 individuals who did not provide age 29 personality, and 29 individuals for whom marriage and parent status were not able to be determined. The final sample included 1988 individuals (80.9% of the original sample; 60.2% female, 63.9% MZ), consisting of 905 pairs and 178 singletons. I compared those who remained in the study to those who did not on age 17 personality. Participants included in the sample did not significantly differ on the trait negative emotionality; however, there were small but significant differences on the traits positive emotionality (d = .156) and constraint (d = .212). Mean ages at the age 17 and 29 assessments were 17.79 (SD = .68) and 29.38 (SD = .60), respectively.

Measures

Personality

Personality was assessed with a 198-item version of the Multidimensional Personality Questionnaire (MPQ; Tellegen & Waller, 2008), which is a self-report inventory measuring 11 primary scales from which 3 higher-order factors can be derived (positive emotionality (PEM), negative emotionality (NEM), and constraint (CON)). These super-factors represent individual differences in how people regulate their emotional states and behavior. Items are endorsed on a scale from *1 = Definitely True* to *4 = Definitely False*, and items are scored so that higher scores represent higher levels of the trait. Scores were transformed to the T-score metric ($M = 50$, $SD = 10$) based on the entire sample (including ages 17 and 29). Of primary interest in the current study were the super-factors.

Life Events

Information regarding marriage and childbirth was obtained at ages 20, 24, and 29 from an interview developed specifically for use with the MTFS (called the Life Events Interview). The marriage status of an individual was determined by a "yes" response to the item "Have you married during the last x years?" (x representing the lag since they were last assessed) while the parent status of an individual was determined by a "yes" response to the item "Have you given birth to [fathered] a child?" If individuals reported marriage or parenthood at any assessment, they were categorized as "ever married" or "ever parent" while those who did not report the event at any age were considered "never married" or "never parent." Individuals were then categorized into 1 of 4 groups based on

their status as well as that of their twin; in the analysis, twins concordant for event occurrence = "concordant married/parent," twins concordant for event non-occurrence = "concordant never married/parent," twins who experienced the event but whose sibling did not = "discordant married/parent", or twins who did not experience the event but whose sibling did = "discordant never married/parent."

Statistical Analyses

First, at the individual level, I compared personality trait means between those who married and those who did not as well as between those who had children and those who had not with linear mixed models. These models included random effects for both family and individual (since observations were nested within these factors) and fixed effects for sex, assessment (17 or 29), the dichotomous variable (ever or never) regarding the event of interest (marriage or parenthood), and their interactions. These models addressed my first question regarding the magnitude of change that occurs across traits from age 17 to 29 as well as my second question regarding the direction of influence between personality and life events. That is, the extent to which personality changes across this developmental period is reflected in the estimate of the effect of assessment on the phenotype. Additionally, the extent to which marriage or parenthood affects personality change from age 17 to 29 is reflected in the estimate of the interaction between group membership and assessment.

To address my last question, I evaluated these associations within twin pairs using the co-twin control method. Specifically, I conducted linear mixed models with family and individual as random effects and fixed effects for sex, assessment (17 or 29),

zygosity, pair-level life event occurrence, and selected interactions. Of interest was whether there were significant within-pair effects of getting married or becoming a parent as well as whether these effects differed by zygosity. Co-twin control analyses were based on assumptions regarding the similarity between twins; specifically, because all twins were reared together, they shared 100% of their rearing environment. However, while MZ twins share 100% of their genetic material, DZ twins share only 50% of their segregating genes. This means that differences in personality between MZ twins must be due to non-shared environmental influences, but differences between DZ twins will be due to both non-shared environmental and genetic influences. The influence of marriage and parenthood was assessed as a non-shared environmental influence in the case of twins discordant for those events. If the effect of an exposure maintains the same magnitude regardless of the relationship between the pairs examined, it is possible that the exposure has a true causal effect; on the other hand, if the magnitude of the exposure effect decreases in related pairs, then there are likely other confounds that explain the exposure effect, and comparison of MZ and DZ twins may indicate to what extent the confounds are genetic or environmental in nature (e.g., if shared environmental confounds entirely explain an exposure effect in unrelated individuals, the effect should disappear for both discordant MZs and DZs; however, if genetic confounds explain an exposure effect, then the effect should disappear for discordant MZs but not entirely for DZs).

Results

Of the 1988 twins included in the sample, 1154 (58.2%) reported getting married and 869 (44.4%) reported having children at some point between ages 17 and 29. Furthermore, 579 pairs were concordant for marriage (322 discordant pairs), and 593 pairs were concordant for parenthood (296 discordant pairs); data were not available for 178 of the twins' siblings. Tables 8 and 9 present the frequencies of individuals in each of the 4 event groups by sex, assessment, and zygosity.

Table 10 shows mean T-scores across MPQ super-factors by sex and assessment. Means suggest sex differences as well as age (assessment) differences in constraint and negative emotionality. Results from the mixed models corroborate these differences, as shown below.

Individual-level Analyses

Tables 11 and 13 present results of tests for fixed effects in the individual-level models for marriage and parenthood, respectively. In Tables 12 and 14, the intercept represents the mean personality T-score at age 17 for males who never experienced the event (marriage or parenthood). Estimates of fixed effects in these mixed models represent the difference between the reference group and other groups in T-score points. Evidence for a relationship between personality change and role transitions would be suggested by observing a significant event by assessment interaction because it would indicate that those who experienced the event changed to a different degree than those who did not.

Marriage. Individual-level effects of marriage are presented in Table 12 (Table 11 displays tests of these fixed effects). In these analyses, I observed a sex effect of about

half of a standard deviation (β_{female} = 4.83) for CON but much smaller effects for NEM (β_{female} = -1.60) and PEM (β_{female} = -0.92). Additionally, estimates suggested an increase in CON of almost a half of a standard deviation (β_{29} = 4.21) as well as an even larger decrease in NEM (β_{29} = -5.96) by age 29. There did not appear to be an age effect on PEM. There was a small but significant marriage group effect on CON (β_{marry} = 1.03), but there was no marriage group effect on NEM or PEM. However, I did observe small but significant interactions between ever marrying and age (assessment) for both CON ($\beta_{marry*29}$ = 2.17) and NEM ($\beta_{marry*29}$ = -2.22).

Parenthood. Individual-level effects of parenthood are presented in Table 14 (Table 13 displays tests of these fixed effects). In these analyses, I observed a sex effect of about half of a standard deviation (β_{female} = 5.19) for CON but much smaller effects for NEM (β_{female} = -2.05) and PEM (β_{female} = -0.19). Additionally, estimates suggested an increase in CON of almost a half of a standard deviation (β_{29} = 4.26) as well as an even larger decrease in NEM (β_{29} = -7.12) by age 29. There did not appear to be an age effect on PEM. I observed no significant parenthood group effect on any super-factor. However, I did observe a significant interaction between parent status and age (assessment) for CON ($\beta_{parent*29}$ = 2.80).

I did not observe interactions between sex and any other variables, with exception of a possible small sex by age effect on PEM ($\beta_{29*female}$ = -1.55). Thus, I did not consider sex interactions in the pair-level analyses.

Pair-level Analyses

Pair-level effects of marriage and parenthood are displayed in Tables 16 and 18, in which the intercept represents the mean personality T-score at age 17 for MZ males who never experienced the event (marriage or parenthood) but whose twin did. Estimates of fixed effects in these mixed models represent the difference between the reference group and other groups in T-score points. The corresponding tests of fixed effects can be found in Tables 15 and 17. The major focus of the co-twin control method is the comparison between twins discordant for the event; as such, the effects of being in the group of discordant twins who did marry or become a parent are of primary interest.

Marriage. Pair-level effects of marriage are presented in Table 16 (Table 15 displays tests of fixed effects). In these analyses, I observed a sex effect of about half of a standard deviation ($\beta_{female} = 5.07$) for CON but much smaller effects for NEM ($\beta_{female} = -1.96$) and PEM ($\beta_{female} = -1.40$). Additionally, estimates suggested an increase in CON of over a half of a standard deviation ($\beta_{29} = 6.32$) as well as a decrease in NEM ($\beta_{29} = -6.83$) by age 29. There did not appear to be an age effect on PEM. Furthermore, there were no significant effects of zygosity on any super-factor. Marriage-discordant twins were not significantly different in any personality trait; additionally, I did not observe significant pair-level interactions between age (assessment) and marriage. Three-way interactions among marriage group, age, and zygosity were not significant for any super-factor, but estimates were substantially larger than zero.

Figures 5-7 show the model-estimated means across marriage groups at ages 17 and 29 by zygosity. MZ twins discordant for marriage were no different in CON (Figure 5) at age 17 ($M_{never} = 46.24$ and $M_{ever} = 46.83$), and both twins' CON scores increased by

age 29 by about the same degree (M_{never} = 52.56 and M_{ever} = 53.63). Similarly, DZ twins discordant for marriage were no different in CON at age 17 (M_{never} = 45.36 and M_{ever} = 45.93); however, it does appear that those who married increased more in CON by age 29 than did those who did not marry (change from age 17 to 29: M_{never} = 4.90 and M_{ever} = 7.10), even though the interaction with zygosity was not significant. Marriage-discordant MZs were also not very different in NEM (Figure 6) at age 17 (M_{never} = 53.78 and M_{ever} = 53.41) or at age 29 (M_{never} = 46.95 and M_{ever} = 45.53). Marriage-discordant DZs, on the other hand, did seem to change differently; even though their NEM scores were similar at age 17 (M_{never} = 52.91 and M_{ever} = 53.47), it does appear that those who married decreased more in NEM by age 29 than those who did not marry (change from age 17 to 29: M_{never} = -4.79 and M_{never} = -8.29), though the interaction with zygosity was not significant. As previously mentioned, PEM did not change much over time for any group (Figure 7).

Parenthood. Pair-level effects of parenthood are presented in Table 18 (Table 17 displays tests of fixed effects). In these analyses, I observed a sex effect of about half of a standard deviation (β_{female} = 5.08) for CON but much smaller effects for NEM (β_{female} = -1.93) and PEM (β_{female} = -1.34). Additionally, estimates suggested an increase in CON of over a half of a standard deviation (β_{29} = 5.61) as well as a large decrease in NEM (β_{29} = -7.60) by age 29. There did not appear to be an age effect on PEM. Furthermore, there were no significant effects of zygosity on any super-factor. Parenthood-discordant twins were not significantly different in any personality trait; additionally, I did not observe significant pair-level interactions between age (assessment) and parenthood. Three-way

interactions among parenthood group, age, and zygosity were not significant for any super-factor, but estimates were substantially larger than zero.

Figures 8-10 show the model-estimated means across parenthood groups at ages 17 and 29 by zygosity. MZ twins discordant for parenthood were no different in CON (Figure 8) at age 17 (M_{never} = 47.06 and M_{ever} = 47.22), and both twins' CON scores increased by age 29 by about the same degree (M_{never} = 52.68 and M_{ever} = 54.01). Similarly, DZ twins discordant for marriage were no different in CON at age 17 (M_{never} = 46.94 and M_{ever} = 46.06); however, it does appear that those who married increased more in CON by age 29 than did those who did not marry (change from age 17 to 29: M_{never} = 4.47 and M_{ever} = 8.35), even though the interaction with zygosity was not significant. Marriage-discordant MZs were also not very different in NEM (Figure 9) at age 17 (M_{never} = 53.68 and M_{ever} = 53.75) or at age 29 (M_{never} = 46.08 and M_{ever} = 46.55). Marriage-discordant DZs, on the other hand, did seem to change differently; their NEM scores were somewhat different at age 17 (M_{never} = 53.30 and M_{ever} = 55.30), which may explain why it appears that those who married decreased more in NEM by age 29 than those who did not marry (change from age 17 to 29: M_{never} = -6.02 and M_{never} = -8.17), though the interaction with zygosity was not significant. As previously mentioned, PEM did not change over time for any group (Figure 10).

Discussion

I examined the change in personality that occurs from late adolescence to early adulthood and influences of entering into marriage and parenthood on that change. Specifically, I compared change in three domains of personality, constraint, negative

emotionality, and positive emotionality, between those individuals who had married and/or become parents and those who had not between ages 17 and 29. Furthermore, I assessed the extent to which any relationships between these life events and personality change were causal or due to shared familial factors by incorporating a within-pair analysis of both MZ and DZ twins. I observed the following patterns. First, a significant amount of change occurred from age 17 to 29 in the CON and NEM super-factors, whereas PEM appeared to be stable over time, and these patterns did not differ between sexes. Second, there was a significant relationship between personality maturity and life events such that individuals who married increased more in CON and decreased more in NEM than those who did not, and those who had children increased more in CON than those who did not. Third, co-twin control analyses suggested that these relationships were not causal, given that these patterns were not observed between identical twins discordant for marriage and parenthood.

The maturity principle proposes that personality develops in a way that allows individuals to function as productive members of society (Caspi et al., 2005). This principle predicts that we should observe change in those characteristics that relate to successfully taking on adult social roles; such characteristics might include being responsible and reliable, interacting with others in a positive manner, and being emotionally stable. In terms of the MPQ, this translates to an increase in constraint and a decrease in negative emotionality, as the three primary scales of the MPQ that most strongly determine the constraint super-factor are control (the extent to which one is planful and sensible), harm avoidance (the opposite of excitement-seeking), and

traditionalism (the extent to which one endorses high moral standards), and the three that most strongly determine negative emotionality are stress reaction (the extent to which one feels anxious and vulnerable), alienation (the extent to which one feels mistreated), and aggression (the extent to which one is physically aggressive and vindictive). In the current study, our findings did in fact support the maturity principle. In both males and females, CON increased by about half of a standard deviation while NEM decreased by nearly three quarters of a standard deviation. This is generally in line with past research; findings from a large meta-analysis (Roberts et al., 2006) and also from recent, large, cross-sectional (Allemand, Zimprich, & Hendriks, 2008; Soto, John, Gosling, & Potter, 2011) and longitudinal studies (Bleidorn, Kandler, Riemann, Angleitner, & Spinath, 2009; Donnellan, Conger, & Burzette, 2007; Hopwood et al., 2011; Specht et al., 2011 ; Vaidya, Gray, Haig, Mroczek, & Watson, 2008) suggest an increase in conscientiousness and agreeableness as well as a decrease in neuroticism. The fact that the effect sizes were in some cases different from those in the current study can most likely be attributed to the differences in personality models at lower levels of trait specificity. For example, aggression is reflected in the NEM dimension in the MPQ whereas it is reflected in the agreeableness (reversed) dimension of the Big Five.

Social Investment Theory proposes that personality will mature in response to social role transitions like marrying and becoming a parent. In the individual-level analyses, I observed that marriage did in fact predict maturity in terms of increasing CON and decreasing NEM relative to unmarried peers; unfortunately, availability of studies examining effects of marriage specifically is limited, as most research has focused

primarily on event composites. Our findings do agree with research showing a desistence from antisocial behavior in men after marriage (Burt et al., 2010). On the other hand, while Specht et al. (2011) did observe decreased extraversion and openness and Anusic, Yap, and Lucas (2014) observed increased life satisfaction after marriage, they did not observe the expected decrease in neuroticism or negative affect (respectively) or increase in conscientiousness and agreeableness. It is reassuring, though, that some studies have presented evidence that beginning one's first substantial romantic relationship affects personality development in the expected direction. That is, researchers have observed decreases in neuroticism, anxiety, depression, impulsivity, and shyness as well as increases in extraversion, self-esteem, life satisfaction, and conscientiousness in those individuals who entered into romantic relationships relative to those who did not enter into such relationships (Lehnart, Neyer, & Eccles, 2010; Neyer & Lehnart, 2007; Wagner, Becker, Ludtke, & Trautwein, 2015).

In our individual-level analyses, I also observed an increase in CON in those young adults who became parents relative to those who did not, which supports SIT; contrastingly, Specht and colleagues (2011) observed a decrease in conscientiousness and Jokela, Kivimaki, Elovainio, and Keltikangas-Jarvinen (2009) reported an increase in emotionality (neuroticism) after having children, which is actually the opposite of what SIT would predict. However, some recent findings may be key in understanding what is going on here. Hutteman and colleagues (Hutteman, Bleidorn, Kerestes, Brkovic, Butkovic, & Denissen, 2014) observed that young adults who reported experiencing significant stress in response to becoming parents decreased in agreeableness and

conscientiousness and increased in neuroticism. This suggests that the way the social transition is viewed and experienced may be an important moderator of personality change, and it is likely not limited to parenthood. For example, researchers have observed effects of relationship quality on personality change, such that being in a relationship with high conflict resulted in increased Neuroticism and decreased Agreeableness (Parker, Ludtke, Trautwein, & Roberts, 2012; Robins, Caspi, & Moffitt, 2002) in early adulthood. Such findings regarding the impact of marriage and parenthood on individual change in personality support SIT but do not provide evidence of a causal effect.

Developmental personality psychologists have proposed that the best test of SIT would be a prospective co-twin control study of social role transitions and personality (e.g., Bleidorn, Kandler, & Caspi, 2014). In our within-pair analyses, I compared personality change between twins discordant for marriage and parenthood. I did not observe any significant differences in personality change between those twins who experienced the event and those who did not. This suggests that our observations of relationships between marriage/parenthood and personality maturity were not causal in nature but instead driven by familial factors (genes or shared environment), which does not support SIT. While other studies have included relatively similar control (non-event) groups for comparison (through careful selection or propensity score matching) to account for selection effects (Anusic et al., 2014; Jonkmann et al., 2013; Wagner et al., 2015), such studies are unable to rule out the possibility of familial factors contributing to both event occurrence and personality change. To our knowledge, only a few studies have investigated the causal effects of life events on personality-relevant phenotypes, but

those that have employed the co-twin control method have successfully uncovered some significant results. Burt et al. (2010) found that men who married were more likely to discontinue antisocial behavior than those who did not marry, and that this relationship was in fact causal. Likewise, Middeldorp and colleagues (2008) observed a causal effect of negative life events (e.g., death of a family member, divorce, assault) on anxious depression and neuroticism. Furthermore, Brown et al. (2014) demonstrated a causal effect of trauma exposure on psychopathology, including major depression, anxiety, and substance abuse. It does not seem to be a fault of the methodology that has prevented the observation of causal effects of social role transitions. Rather, it may be the case that causal effects of events are only observed when the phenotype of interest is dysfunctional in nature.

It is also plausible, of course, that it is only marriage and parenthood that are not particularly powerful instigators for personality maturation while other social role transitions do in fact provide significant provocation for change. Future research should investigate the causal effects of work-relevant milestones on personality maturity, as many researchers have observed significant relationships between personality development and starting a job, unemployment, and retirement (Leikas & Salmela-Aro, 2015; Specht et al., 2011).

Advantages to the current study include the narrow age range, which allowed us to focus on a specific developmental period, measurement of events that were not dependent on subjective assessments but rather verifiable occurrences, and the inclusion of identical twins to examine causal effects of social role transitions. However, there are

some limitations that must be considered. First, as in most personality studies, there may be concerns about the validity of self-reports of personality. However, validity of self-reports has been demonstrated in that they correlate highly with observer-reports and that they predict personality-relevant criterion variables (Connelly & Ones, 2010). Second, though the overall sample was large, the cell sizes in the within-pair analyses were much more modest (< 100 in the majority of cells), resulting in larger standard errors. While estimates of the interaction effects of group, assessment, and zygosity were greater than zero, they are not significant in comparison to their standard errors. This makes it difficult to determine whether or not shared familial factors explaining the relationship between personality development and role transitions are environmental or genetic in nature. Third, I considered individuals to be "ever married" if they reported getting married at any assessment, without considering the length of marriage or whether they were still married. It is reasonable to assume that an individual who marries and then separates within a short period of time has probably not invested enough for it to constitute a social role transition. Additionally, individuals were considered "ever parents" if they reported ever giving birth or fathering a child. This disregards the possibility of some individuals adopting children as well as others giving up their children for adoption. If it is investment in the social role of parenthood that is important for personality development, then adopting children should have the same effects as raising biological children while simply giving birth but not raising a child should not have those effects.

It is also possible that effects of social role transitions may not be observed at the domain level of personality but still have some effect on more specific trait variation. Research has shown that differences exist in the underlying etiology of facet-level versus domain-level traits. Variance at the domain level is highly heritable and not significantly influenced by shared environmental influences while variance at the facet level is also heritable but sometimes also differentially influenced by environmental factors (Jang, McCrae, Angleitner, Riemann, & Livesley, 1998; Kandler, Riemann, Spinath, & Angleitner, 2010). Future research should explore the extent to which social role transitions can explain environmental influence on facet-level traits.

Finally, I did not necessarily examine *investment* in social roles, since I didn't examine any moderators of the relationship between events and personality change (such as the circumstances under which individuals married or had kids). However, as Neyer and colleagues (2014) mentioned, being married to some extent assumes an investment in the relationship in order to maintain that marriage. As for parenting, it seems that having children would require a substantial investment even if a parent did not necessarily want to invest (e.g., legal obligations).

Table 8. Number of participants with personality data across marriage group, time, sex, and zygosity.

	Discordant Never Marry	**Discordant Ever Marry**	**Concordant Never Marry**	**Concordant Ever Marry**	**Totals**
Male MZ	67	67	130	200	464
Male DZ	50	50	64	66	230
Female MZ	118	118	152	300	688
Female DZ	87	87	84	162	420
Totals	322	322	430	728	1802

Table 9. Number of participants with personality data across parenthood group, time, sex, and zygosity.

	Discordant Never Parent	**Discordant Ever Parent**	**Concordant Never Parent**	**Concordant Ever Parent**	**Totals**
Male MZ	76	76	188	116	456
Male DZ	42	42	96	44	224
Female MZ	101	101	252	228	682
Female DZ	77	77	164	98	416
Totals	296	296	700	486	1778

Table 10. Mean (SD) personality trait T-scores by sex and age.

		CON	NEM	PEM
Male				
	Age 17	44.15 (9.24)	54.80 (9.22)	50.79 (9.43)
	Age 29	49.48 (9.27)	47.62 (9.22)	51.14 (9.81)
Female				
	Age 17	49.04 (9.47)	52.69 (9.65)	50.04 (10.58)
	Age 29	55.13 (8.92)	45.77 (9.10)	48.69 (9.76)

Table 11. Tests of fixed effects from linear mixed models considering marriage effects at the individual level.

		CON	NEM	PEM
Marriage		1, 1910.8	1, 1914.6	1, 1909.1
df		32.48	9.06	1.37
	F	< .001	.003	.24
	p			
Age		1, 1928.8	1, 1945.0	1, 1946.6
df		766.13	1037.35	5.05
	F	< .001	< .001	.025
	p			
Sex		1, 1095.2	1, 1061.8	1, 1091.7
df		141.31	17.76	11.71
	F	< .001	< .001	< .001
	p			
Marriage*Age	**df**	1, 1928.8	1945.0	1, 1946.6
	F	22.17	20.15	0.21
	p	< .001	< .001	.65
Marriage*Sex		1, 1910.8	1, 1914.6	1, 1909.1
df		0.02	0.60	0.21
	F	0.88	.44	.65
	p			
Age*Sex		1, 1928.8	1, 1945.0	1, 1946.6
df		2.41	0.53	13.16
	F	0.12	.47	< .001
	P			
Marriage*Age*Sex	**df**	1, 1928.8	1, 1945.0	1, 1946.6
	F	0.41	0.47	1.28
	p	0.52	.49	.26

Note. Tests are based on Satterthwaite approximation for degrees of freedom.

Table 12. Parameter estimates from linear mixed models considering marriage effects at the individual level.

	CON	**NEM**	**PEM**
Random Effects (SD)			
Within-Person	4.19	3.92	4.39
Within-Family	5.36	5.46	6.00
Residual	6.09	6.45	6.59
Fixed Effects estimate (SE)			
Intercept	43.49 (.51)	54.74 (.52)	50.59 (.55)
Ever Married	1.03 (.64)	0.30 (.65)	0.49 (.69)
Age 29	4.21 (.46)	-5.96 (.49)	0.15 (.50)
Female	4.83 (.68)	-1.60 (.69)	-0.92 (.73)
Ever Married*Age 29	2.17 (.62)	-2.22 (.66)	0.30 (.68)
Ever Married*Female	0.15 (.83)	-0.86 (.85)	0.14 (.90)
Age 29*Female	0.89 (.61)	0.02 (.65)	-1.10 (.66)
Ever Married*Female*29	-0.52 (.81)	0.59 (.86)	-0.99 (.88)

Note. Estimates are presented in the T-score metric. Reference group = male, 17, never married twins.

Table 13. Tests of fixed effects from linear mixed models considering parenthood effects at the individual level.

		CON	NEM	PEM
Parenthood	**df**	1, 1900.2	1, 1906.7	1, 1897.1
	F	17.81	2.68	0.08
	p	< .001	.10	.77
Age	**df**	1, 1907.6	1, 1921.8	1, 1923.0
	F	852.96	1070.17	4.47
	p	< .001	< .001	.035
Sex	**df**	1, 1084.0	1, 1051.5	1, 1079.8
	F	135.47	17.64	13.10
	p	< .001	< .001	< .001
Parenthood*Age	**df**	1, 1907.6	1, 1921.8	1, 1923.0
	F	40.57	0.79	0.65
	p	< .001	.38	.42
Parenthood*Sex	**df**	1, 1900.2	1, 1906.7	1, 1897.1
	F	1.08	0.02	4.02
	p	.30	.88	.045
Age*Sex	**df**	1, 1907.6	1, 1921.8	1, 1923.0
	F	1.81	0.26	15.79
	P	.18	.61	< .001
Parenthood*Age*Sex	**df**	1, 1907.6	1, 1921.8	1, 1923.0
	F	0.28	0.31	0.22
	p	.60	.58	.64

Note. Tests are based on Satterthwaite approximation for degrees of freedom.

Table 14. Parameter estimates from linear mixed models considering parenthood effects at the individual level.

	CON	**NEM**	**PEM**
Random Effects (SD)			
Within-Person	4.24	3.93	4.33
Within-Family	5.42	5.46	6.01
Residual	6.06	6.46	6.61
Fixed Effects estimate (SE)			
Intercept	43.83 (.46)	54.49 (.47)	50.65 (.50)
Ever Parent	0.54 (.65)	0.74 (.67)	0.40 (.70)
Age 29	4.26 (.41)	-7.12 (.43)	0.13 (.44)
Female	5.19 (.61)	-2.05 (.62)	-0.19 (.66)
Ever Parent*Age 29	2.80 (.63)	-0.14 (.68)	0.56 (.69)
Ever Parent*Female	-0.55 (.84)	0.12 (.86)	-1.37 (.91)
Age 29*Female	0.76 (.53)	0.46 (.57)	-1.55 (.58)
Ever Parent*Female*29	-0.43 (.81)	-0.48 (.87)	-0.41 (.89)

Note. Estimates are presented in the T-score metric. Reference group = male, 17, never parenting twins.

Table 15. Tests of fixed effects from linear mixed models considering marriage effects at the pair level.

		CON	NEM	PEM
Marriage	**df**	3, 897.05	3, 894.37	3, 897.49
	F	11.06	3.05	1.62
	p	< .001	.028	.18
Age	**df**	1, 1761.08	1, 1771.67	1, 1774.43
	F	670.16	861.80	3.56
	p	< .001	< .001	.059
Sex	**df**	1, 894.67	1, 889.69	1, 893.24
	F	114.50	17.63	7.27
	p	< .001	< .001	.007
Zygosity	**df**	1, 897.01	1, 892.18	1, 895.52
	F	0.97	0.35	0.03
	p	.33	.56	.87
Marriage*Age	**df**	3, 1760.65	3, 1771.33	3, 1774.10
	F	9.84	9.12	0.89
	p	< .001	< .001	.45
Marriage*Age*Zygosity	**df**	3, 1760.75	3, 1771.33	3, 1774.10
	F	1.49	3.04	0.38
	p	.22	.028	.77

Note. Tests are based on Satterthwaite approximation for degrees of freedom.

Table 16. Parameter estimates from linear mixed models considering marriage effects at the pair level.

	CON	**NEM**	**PEM**
Random Effects (SD)			
Within-Person	4.18	3.90	4.36
Within-Family	5.41	5.28	6.02
Residual	6.11	6.38	6.62
Fixed Effects estimate (SE)			
Intercept	43.13 (.74)	54.98 (.74)	51.42 (.81)
DIS Ever Married	0.59 (.77)	-0.36 (.78)	-0.94 (.83)
CON Never Married	0.42 (.93)	-0.35 (.93)	-1.12 (1.01)
CON Ever Married	2.10 (.83)	-1.15 (.83)	0.48 (.90)
Female	5.07 (.47)	-1.96 (.47)	-1.40 (.52)
Age 29	6.32 (.64)	-6.83 (.67)	-0.91 (.69)
DZ	-0.88 (1.04)	-0.87 (1.04)	-1.14 (1.13)
DIS Ever Married*age29	0.48 (.91)	-1.05 (.95)	0.50 (.98)
CON Never Married*age29	-2.21 (.82)	0.79 (.86)	0.79 (.89)
CON Ever Married*age29	-0.12 (.75)	-0.48 (.78)	-0.22 (.81)
DIS Ever Married*age29*DZ	1.73 (1.40)	-2.45 (1.46)	-1.29 (1.51)
CON Never Married*age29*DZ	1.13 (1.33)	-1.12 (1.38)	-1.39 (1.44)
CON Ever Married*age29*DZ	2.47 (1.21)	-3.50 (1.26)	-1.11 (1.31)

Note. Estimates are presented in the T-score metric. Reference group = male, 17, MZ, never married twins with married sibling. DIS = discordant twins, CON = concordant twins.

Table 17. Tests of fixed effects from linear mixed models considering parenthood effects at the pair level.

		CON	NEM	PEM
Parenthood	**df**	3, 884.52	3, 881.54	3, 884.57
	F	4.94	2.11	0.58
	p	.002	.10	.63
Age	**df**	1, 1736.60	1, 1747.65	1, 1750.38
	F	708.93	868.46	8.64
	p	< .001	< .001	.003
Sex	**df**	1, 881.83	1, 876.30	1, 880.32
	F	110.86	16.78	6.51
	p	< .001	< .001	.011
Zygosity	**df**	1, 884.50	1, 879.21	1, 882.97
	F	0.35	0.81	0.36
	p	.56	.37	.55
Parenthood*Age	**df**	3, 1736.64	3, 1747.70	3, 1750.42
	F	14.90	0.80	2.18
	p	< .001	.50	.09
Parenthood*Age*Zygosity	**df**	3, 1736.63	3, 1747.69	3, 1750.42
	F	1.43	1.21	1.43
	p	.23	.30	.23

Note. Tests are based on Satterthwaite approximation for degrees of freedom.

Table 18. Parameter estimates from linear mixed models considering parenthood effects at the pair level.

	CON	**NEM**	**PEM**
Random Effects (SD)			
Within-Person	4.23	3.89	4.32
Within-Family	5.49	5.29	6.05
Residual	6.10	6.42	6.63
Fixed Effects estimate (SE)			
Intercept	43.93 (.75)	54.87 (.74)	51.72 (.81)
DIS Ever Parent	0.15 (.79)	0.07 (.80)	-0.16 (.84)
CON Never Parent	0.26 (.86)	-1.11 (.86)	-0.67 (.93)
CON Ever Parent	0.41 (.91)	-0.41 (.90)	-0.73 (.98)
Female	5.08 (.48)	-1.93 (.47)	-1.34 (.52)
Age 29	5.61 (.65)	-7.60 (.68)	-0.97 (.71)
DZ	-0.12 (1.10)	-0.38 (1.10)	-0.24 (1.19)
DIS Ever Parent*age29	1.18 (.92)	0.40 (.97)	-0.47 (1.01)
CON Never Parent*age29	-1.11 (.77)	0.79 (.81)	0.33 (.84)
CON Ever Parent*age29	1.42 (.80)	0.76 (.84)	0.55 (.87)
DIS Ever Parent*age29*DZ	2.71 (1.46)	-2.55 (1.54)	2.86 (1.59)
CON Never Parent*age29*DZ	0.95 (1.24)	-1.47 (1.30)	2.54 (1.35)
CON Ever Parent*age29*DZ	1.98 (1.36)	-2.36 (1.43)	2.08 (1.48)

Note. Estimates are presented in the T-score metric. Reference group = male, 17, MZ, never parenting twins with parenting sibling. DIS = discordant twins, CON = concordant twins.

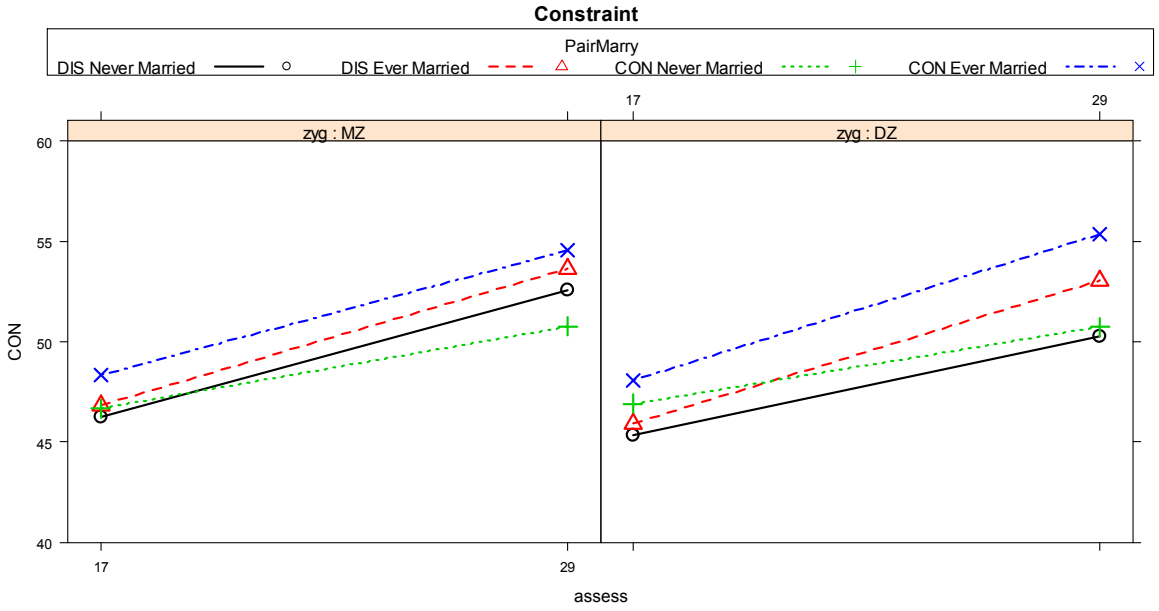

Figure 5. Mean differences in Constraint at ages 17 and 29 across pair-level marriage groups, separated by zygosity.

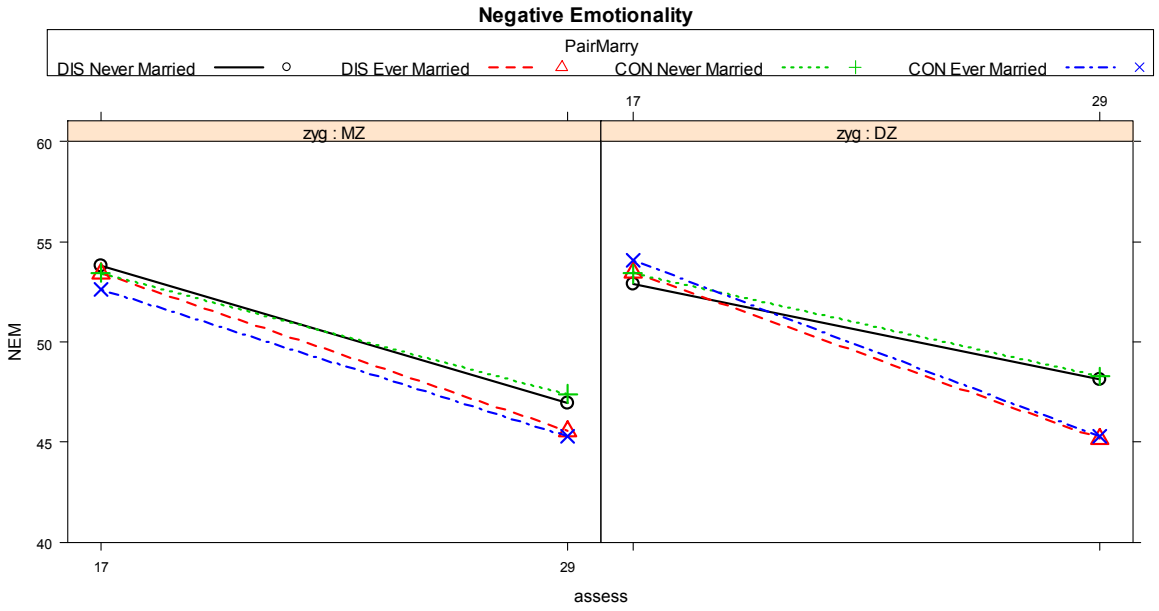

Figure 6. Mean differences in Negative Emotionality at ages 17 and 29 across pair-level marriage groups, separated by zygosity.

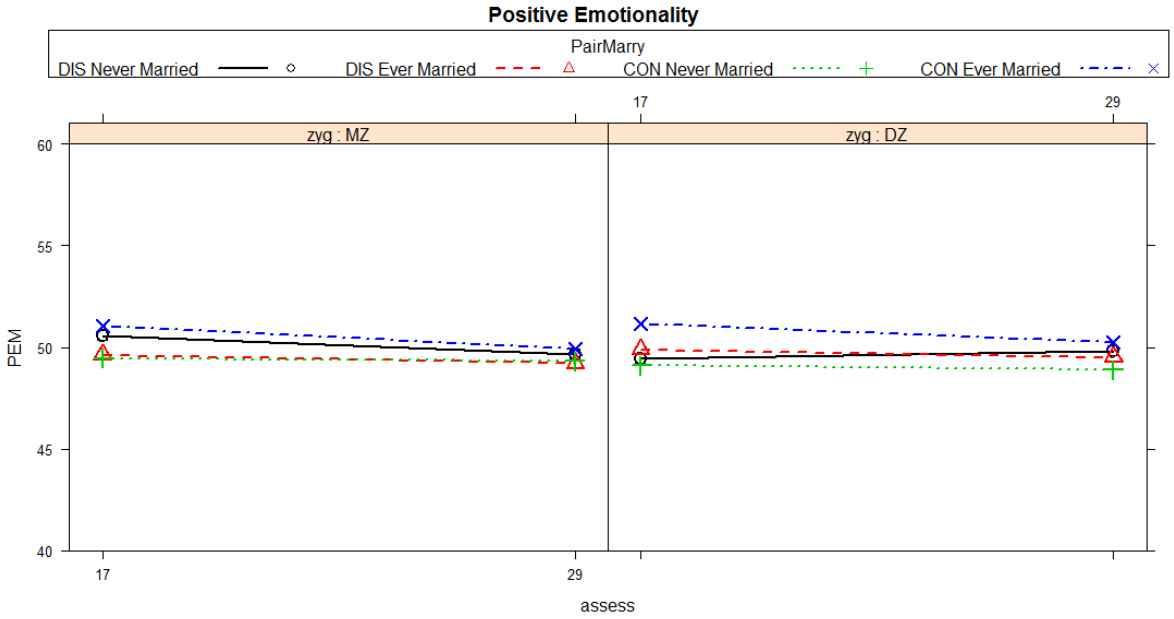

Figure 7. Mean differences in Positive Emotionality at ages 17 and 29 across pair-level marriage groups, separated by zygosity.

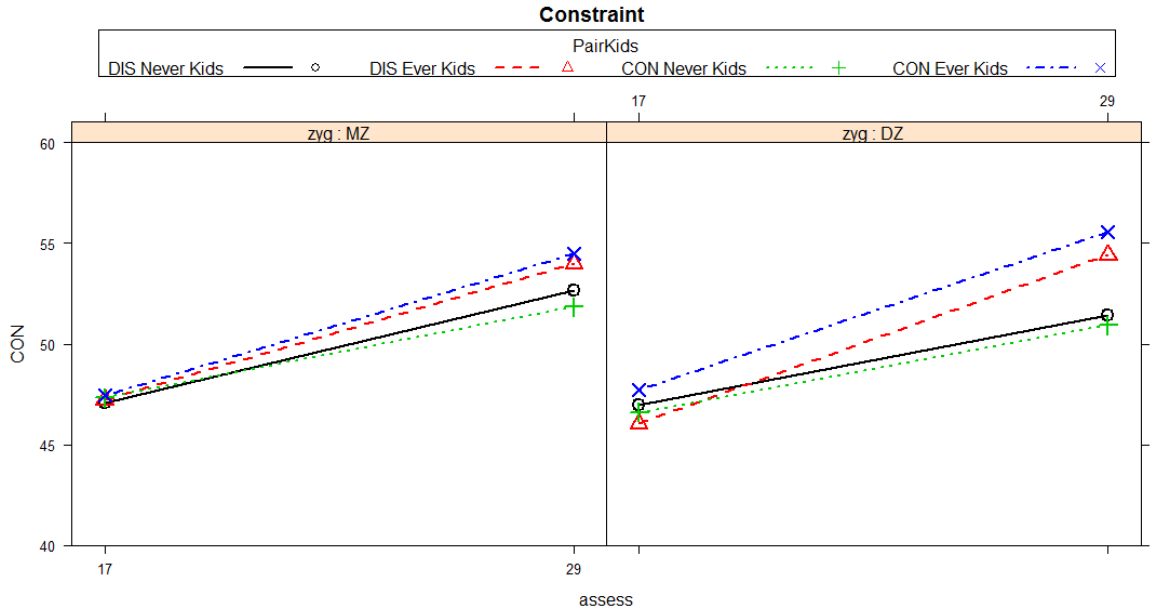

Figure 8. Mean differences in Constraint at ages 17 and 29 across pair-level parent groups, separated by zygosity.

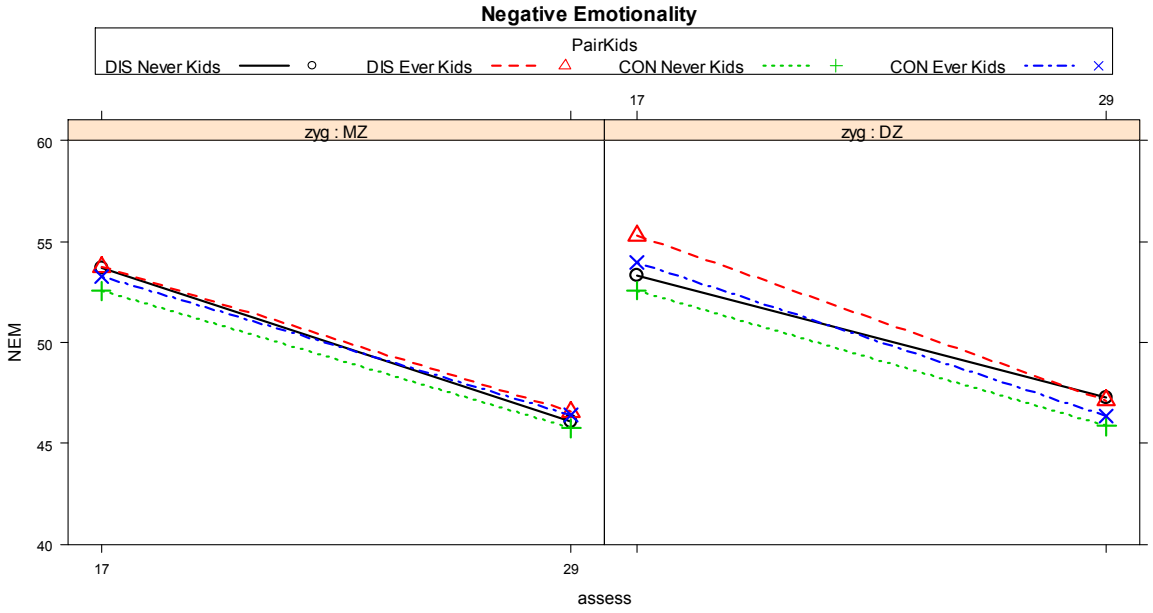

Figure 9. Mean differences in Negative Emotionality at ages 17 and 29 across pair-level parent groups, separated by zygosity.

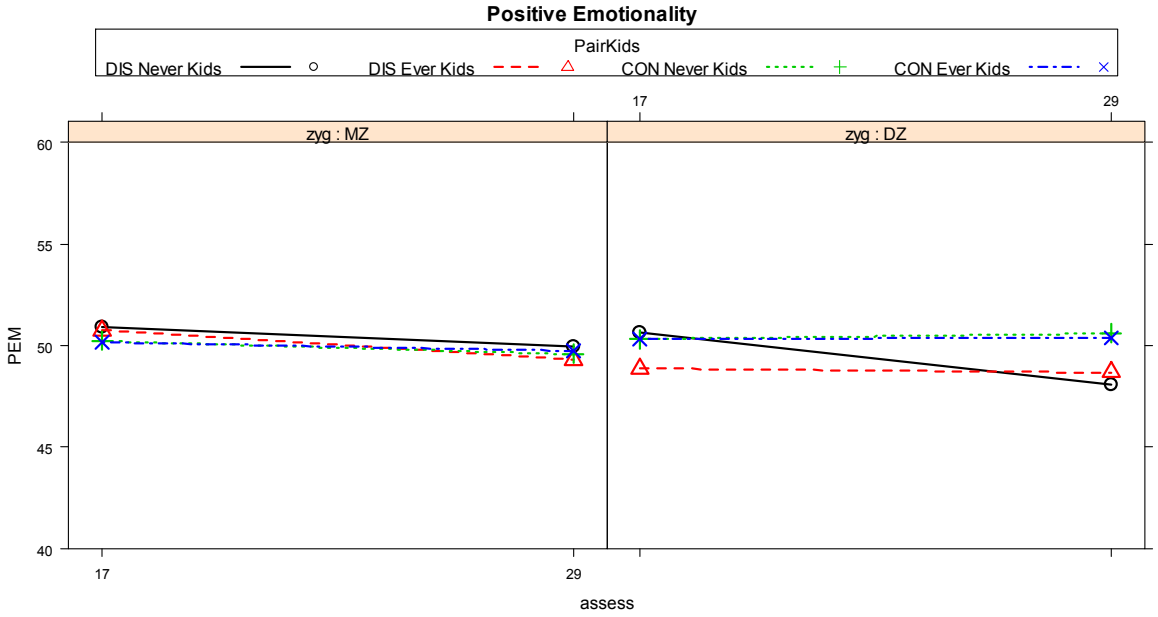

Figure 10. Mean differences in Positive Emotionality across pair-level parent groups.

Chapter 4. Conclusion

In these complementary studies, I have focused on a specific developmental period extending from adolescence through young adulthood to describe 1) the extent to which personality traits are stable from age 14 to 29, 2) the direction of change in personality over these years, 3) the extent to which additive genetic variation (based on common genotyped variants) can explain personality variation at different ages, and 4) whether sufficient evidence exists for a *causal* effect of specific environmental factors (i.e., social role transitions) on personality development through this period. My goals were to investigate opposing theories regarding personality development (FFT and SIT) and to examine whether narrowing personality phenotypes to specific ages would aid in the search for missing heritability and environmentality of personality.

Caspi et al. (2005) proposed personality principles based on what has been observed about personality development in numerous studies. The corresponsive principle describes the fact that, while substantial personality stability is observed across all ages, traits become even more stable as individuals age. Rank-order correlations represent the extent to which individuals are stable in their relative standing on the trait; that is, if a person scores very low (relative to others) in some trait at one measurement, he is likely to still be very low in that trait at a later time, even if actual behavior has changed to some degree. The corresponsive principle specifically refers to the fact that rank-order correlations increase as the sample increases in age, which has been observed in multiple studies (see Roberts & DelVecchio, 2000 for a meta-analysis). In the current

studies, I too observed evidence of significant phenotypic stability of personality, as rank order correlations were greater than .4 across all traits between the ages of 14 and 29. Additionally, I observed evidence of the cumulative continuity principle in that rank-order correlations were much stronger between the ages of 24 and 29 (all > .7) than between earlier ages, even when the time interval was smaller (e.g., correlations were all < .7 between ages 14 and 17).

The maturity principle posits that personality changes reflect the need to behave in different ways in order to be productive members of society. Evidence of this principle resides in the consistent observations of increases in conscientiousness, agreeableness, and emotional stability that are typically experienced from adolescence through adulthood (Roberts et al., 2006). In the current studies, I too observed evidence in line with this principle; I observed significant change over time in the super-factors NEM and CON (but not PEM). Specifically, the degree of NEM (most similar to neuroticism in Big Five models of personality) decreased from age 17 to 29 while the degree of CON (most similar to conscientiousness) increased. Furthermore, even though some sex differences did emerge in overall levels of personality traits, changes over time were similar across sexes.

FFT proposes that personality development is fixed and as such environmental factors do not directly contribute to personality stability or change. It implies high stability of personality after age 30, a large degree of consistency in age trends across cultures, and high heritability of personality stability and change. These predictions do appear to be true to some extent. It may not be the case that personality stops changing

completely after age 30, but it is true that the majority of change occurs earlier in life (as evidenced by the cumulative continuity principle and the maturity principle). Also, cross-cultural studies have in fact shown consistency in age-related personality differences (McCrae et al., 1999; 2000), which would suggest that the core of personality is unaffected by culture, a major source of environmental influence.

Research has also shown that personality is substantially heritable, and that rank order stability and change are to some degree influenced by genetic factors in young adulthood, as evidenced by large but imperfect genetic correlations across ages (Briley & Tucker-Drob, 2014). In my own analyses, I observed substantial heritability estimates at each age (h^2 ~.45), and estimates were in line with previous research (including a recent meta-analysis of twin and family studies; Vukasovic & Bratko, 2015). Also, estimates did not substantially change during this developmental period, which would be predicted by FFT. This does disagree with some previous research that observed a decrease in heritability over time (e.g., Kandler et al., 2010), however, it appears that the majority of the decrease in heritability that is observed across the lifespan occurs prior to age 15 and after age 30 (Briley & Tucker-Drob, 2014). Bivariate twin analyses uncovered large genetic correlations, even across 15 years (r_g = .75), suggesting that genetic factors influencing personality variation in adolescence are much the same as those influencing variation in adulthood. However, genetic influence was not entirely stable, as estimates did not reach unity over that time period. Additionally, genetic correlations increased with age; between ages 24 and 29, correlations did approach unity whereas the genetic correlations between ages 14 and 17 were smaller than that, despite a shorter time

interval. This is in line with previous research; in particular, in their meta-analysis Briley and Tucker-Drob (2014) observed a linear increase in genetic stability until age 30, when it approached unity.

Some studies have examined the relative genetic and environmental influence on absolute personality variation, isolating variance associated with stability from that associated with change using latent growth curve modeling; this approach involves the decomposition of variance on intercepts and slopes separately. Although I did not take this approach, evidence from such studies does support FFT to some extent. Hopwood and colleagues (2011) observed substantial additive genetic influence on stable personality variance (as reflected by the intercept) for both negative emotionality (.46) and constraint (.68) from late adolescence to adulthood, but they also observed genetic influence on change in constraint (.50). Similarly, Bleidorn and colleagues (2009) found genetic influence on slope for agreeableness (.53) and conscientiousness (.70).

While the findings discussed thus far seem to support FFT, they are not incompatible with SIT. As previously discussed, personality is substantially heritable at all ages, most phenotypic change occurs in young adulthood (similarly across cultures), and this coincides with instability in genetic (and environmental) influences; as genetic factors stabilize during this period of time, personality also stabilizes. SIT explains change in terms of entering into new social roles, as each role transition requires behavioral adaptation in order to be successful, and these new roles are usually experienced between late adolescence and early adulthood (20s and 30s). This would account for the fact that most individuals experience personality change around the same

age, even across cultures. The fact that genetic influences are substantial and that they stabilize at the same time is explained as being due at least in part to influence of the corresponsive principle (selection effects) because individuals are increasingly able to choose and mold their own environments as they transition from adolescence to adulthood. The extent to which this represents gene-environment correlation is determined by the extent to which genes control exposure to different environments; this is likely to be considerable given that the heritability of personality is substantial. The key difference between the theories in this respect is that FFT proposes that any association between personality development and environmental factors is mediated through genes while SIT proposes that there are causal influences of environmental factors on personality development even after taking into consideration selection effects. Because effects of gene-environment correlation are contained in the genetic component of a variance decomposition (Purcell, 2002), heritability may in fact be overestimated in twin studies that do not explicitly model such correlations.

Thus, it is ambiguous from past research which theory better accounts for personality development. However, there are some findings that support SIT but are problematic for FFT. Heritability decreases after early adulthood, the corollary being that environmental influence increases (Briley & Tucker-Drob, 2014). This could mean that genetic influences directly contribute less to personality differences, that gene-environment correlation decreases as individuals settle into more permanent life situations, that environmental influences directly contribute more to personality variance, or that gene by (non-shared) environment interactions become more important (which are

modeled in the non-shared environmental component in a variance decomposition in twin models, Purcell, 2002). None of these possibilities align with predictions from FFT. Another problem for FFT is that environmental influence on personality becomes more stable as individuals age. This means that the same factors that are important earlier in adulthood remain important later in adulthood, and the extent to which this is true increases over time. FFT does not allow for stability in environmental influence on personality. Finally, in latent growth curve models, environment explained at least half the variance in slope across traits (Bleidorn et al., 2009; Hopwood et al., 2011). This provides evidence that environmental factors are relevant to personality change over time, which is incompatible with FFT. In fact, Kandler and colleagues (2012) observed significant effects of life events on personality scores across several measurements that were independent of selection effects, though they were small (genetic effects of personality on life events were much stronger).

Recently, researchers have emphasized the need to test hypotheses of SIT by including specific measures of the environment in behavior genetic studies (Bleidorn et al., 2014; Specht et al. 2014). Specifically, they have suggested the use of co-twin control studies to account for familial influences shared between personality and environments and to isolate the possible causal effects of environments such as taking on new adult social roles. In my co-twin control analyses, I did not observe causal effects of getting married or having children, two major role changes individuals experience during young adulthood. If these findings are not due to methodological flaws and are replicated in future research, SIT may not be the best explanation for personality development. Does

that mean that FFT better fits the bill? Ultimately, as Specht et al. (2014) concluded, it is unlikely that any one developmental theory is comprehensive enough to explain all aspects of personality development in early adulthood. It is entirely possible that other social role changes do have causal effects, and it is also possible that some changes are largely genetically influenced.

The co-twin control analyses described failed to uncover any of the "missing environmentality" of personality. Similarly, genetic analyses did not account for much of the "missing heritability." It did not appear that restricting GCTA to different developmental groups made any difference in the results. Based on the current findings, the type of genetic variation relevant for NEM and PEM may not be additive in nature—or, at least it is not highly influenced by effects of *common* SNPs included in our genotyping efforts—because estimates were non-significant for these super-factors and their associated primary scales. Interestingly, however, genetic influences on the constraint super-factor as well as harm avoidance, traditionalism, and aggression primary scales were to some extent accounted for by SNPs in the adult sample (in fact, SNP heritability explained about half of the twin heritability). This suggests the continuing need to explore etiology of personality separately across traits, and that further pursuit of constraint-related traits may be successful in identifying particular regions of the genome responsible for personality variation and even further in the future there may come a time when we can elucidate a biological pathway associated with this variation. That is not to say that the same will not be possible for other traits, but given the results from the current GCTA (as well as others), we may need to change our approach to uncovering

such relevant genetic variation. It may be that rare SNPs and/or structural variation is relevant for these other traits, which should be further pursued in genome sequencing studies. Furthermore, it could be that personality has complex etiology that includes a large degree of gene-environment and gene-gene interactions.

A gene by environment interaction (GxE) implies that there is genetic control of sensitivity to the environment (or vice versa). The push to investigate these interactions occurred after the publication of two highly influential papers by Caspi and colleagues; first, they observed an association between a functional polymorphism in the MAO-A gene and increased risk of aggressive behavior that appeared only when there was history of childhood abuse (Caspi et al., 2002). Second, they found an association between the 5-HTTLPR short allele and increased risk of depression that occurred only in those individuals who were exposed to a stressful environment (Caspi et al., 2003). Importantly, the value of these studies is tied to their replicability across populations (Uher, 2013). A rigorous meta-analysis of studies involving the 5-HTTLPR polymorphism and depression (Risch et al., 2009) suggested no interaction with stressful life events despite the original finding. On the other hand, a meta-analysis of studies involving the MAO-A polymorphism and antisocial behavior did suggest an interaction with childhood maltreatment (Kim-Cohen et al., 2006); across studies, the correlation between maltreatment and antisocial behavior was .32 in the low MAO-A activity group but only .12 in the high MAO-A activity group. Continued interest in these and related interactions has spurred a number of new studies that have both supported (e.g., Cicchetti, Rogosch, & Thibodeau, 2012; Fergusson, Boden, Horwood, Miller, &

Kennedy, 2012) and refuted (e.g., Sadeh, Javdani, & Verona, 2013) the original findings. It is likely that gene-environment interactions are important to the development of personality, but reliable detection of their effects may require more methodologically rigorous research. Future research should focus on detection of such effects. It is possible to incorporate GxE into GWAS (GEWIS) to examine possible interactions across of number of genetic loci in the same analysis, but it is not really feasible because of low statistical power to detect effects. Vrieze and colleagues (2012) suggested the use of polygenic risk scores to cut back on the number of statistical tests performed in such an analysis and emphasized the importance of developmental considerations in any gene-finding effort. Gene-gene interactions (non-additive genetic influence) is not detectable in GWAS in a straightforward way; however, as Vrieze et al. (2012) pointed out, what comprises an environment from a gene's point of view could in fact be the effects of other genes. In that sense, traditional GxE methods could be applied by considering the effects of multiple genetic variants rather than a genetic variant and a physical environment.

Overall, it appears that there is strong evidence for genetic influence on personality development, and that such factors work both directly and indirectly (by the selection of preferred environments, i.e., gene-environment correlations). However, there is also evidence for the importance of non-shared environmental influences as well, and these influences may also work directly or indirectly (by moderating genetic influences through gene-environment interactions). I propose that such gene-environment correlations and interactions need to be accounted for in quantitative genetic studies,

molecular studies, and environmental studies. Doing so will allow us a clearer view of what types of influences to expect and how to more effectively search for them.

Bibliography

Allemand, M., Zimprich, D., & Hendriks, A. A. J. (2008). Age differences in five personality domains across the life span. *Developmental Psychology, 44,* 758-770. doi: 10.1037/0012-1649.44.3.758

Anusic, I., Yap, S. C. Y., & Lucas, R. E. (2014). Does personality moderate reaction and adaptation to major life events? Analysis of life satisfaction and affect in an Australian national sample. *Journal of Research in Personality, 51*, 69-77.

Benning, S. D., Patrick, C. J., Hicks, B. M., Blonigen, D. M., & Krueger, R. F. (2003). Factor structure of the psychopathic personality inventory: Validity and implications for clinical assessment. *Psychological Assessment, 3,* 340-350.

Billig, J. P., Hershberger, S. L., Iacono, W. G., & McGue, M. (1996). Life events and personality in late adolescence: Genetic and environmental relations. *Behavior Genetics, 26,* 543-554.

Bleidorn, W., Kandler, C., & Caspi, A. (2014). The behavioural genetics of personality development in adulthood—classic, contemporary, and future trends. *European Journal of Personality, 28,* 244-255. doi: 10.1002/per.1957

Bleidorn, W., Kandler, C., Riemann, R., Angleitner, A., & Spinath, F. M. (2009). Patterns and sources of adult personality development: Growth curve analyses of the NEO PI-R scales in a longitudinal twin study. *Journal of Personality and Social Psychology, 97,* 142-155. doi: 10.1037/a0015434

Blonigen, D. M., Carlson, M. D., Hicks, B. M., Krueger, R. F., & Iacono, W. G. (2008). Stability and change in personality traits from late adolescence to early adulthood: A longitudinal twin study. *Journal of Personality, 76,* 229-266.

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

Brown, R. C., Berenz, E. C., Aggen, S. H., Gardner, C. O., Knudsen, G. P., Reichborn-Kjennerud, T., . . . Amstadter, A. B. (2014). Trauma exposure and axis I psychopathology: A cotwin control analysis in Norwegian young adults. *Psychological Trauma: Theory, Research, Practice, and Policy, 6,* 652-660. doi: 10.1037/a0034326

Burt, S. A., Donnellan, B., Humbad, M. N., Hicks, B. M., McGue, M., & Iacono, W. G. (2010). Does marriage inhibit antisocial behavior? An examination of selection vs. causation via a longitudinal twin design. *Archives of General Psychiatry, 67,* 1309-1315.

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

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., . . . Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. *Science, 297,* 851-854.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I., Harrington, H. L., . . . Poulton, R. (2003). Influence of life stress on depression: Moderation by polymorphism in the 5-HTT gene. *Science, 301,* 386-389.

Cicchetti, D., Rogosch, F. A., & Thibodeau, E. L. (2012). The effects of child maltreatment on early signs of antisocial behavior: Genetic moderation by tryptophan hydroxylase, serotonin transporter, and monoamine oxidase A genes. *Development and Psychopathology, 24,* 907-928. doi: 10.1017/s0954579412000442

Connelly, B., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin, 136,* 1092-1122. doi: 10.1037/a0021212

Davies, G., Tenesa, A., Payton, A., Yang, J., Harris, S. E., Liewald, D., . . . Deary, I. J. (2011). Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Molecular Psychiatry, 10,* 996-1005 doi: 10.1038/mp.2011.85

de Moor, M. H. M., Costa, P. T., Terracciano, A., Krueger, R. F., de Geus, E. J. C., Toshiko, T., . . . Boomsma, D. I. (2012). Meta-analysis of genome-wide association studies for personality. *Molecular Psychiatry, 17,* 337-349. doi: 10.1038/mp.2010.128

de Moor, M. H. M., van den Berg, S. M., Verweij, K. J. H., Krueger, R. F., Luciano, M., Vasquez, A. A., Matteson, L. K., . . . Boomsma, D. I. (2015). Genome-wide association study identifies novel locus for neuroticism and shows polygenic association with Major Depressive Disorder. *Journal of the American Medical Association Psychiatry, 72,* 642-650. doi: 10.1001/jamapsychiatry.2015.0554

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

Eaves, L., Heath, A., Martin, N., Maes, H., Neale, M., Kendler, K., Kirk, K., & Corey, L. (1999). Comparing the biological and cultural inheritance of personality and social attitudes in the Virginia 30,000 study of twins and their relatives. *Twin Research, 2,* 62-80.

Felson, J. (2014). What can we learn from twin studies? A comprehensive evaluation of the equal environments assumption. *Social Science Research, 43,* 184-189.

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.

Fergusson, D. M., Boden, J. M., Horwood, L. J., Miller, A., & Kennedy, M. A. (2012). Moderating role of the MAOA genotype in antisocial behavior. *British Journal of Psychiatry, 200*, 116-123. doi: 10.1192/bjp.bp.111.093328

Finkel, D., & McGue, M. (1997). Sex differences and nonadditivity in heritability of the Multidimensional Personality Questionnaire scales. *Journal of Personality and Social Psychology, 72,* 929-938.

Fraley, 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.

Heiman, N., Stallings, M. C., Hofer, S. M., & Hewitt, J. K. (2003). Investigating age differences in the genetic and environmental structure of the tridimensional

personality questionnaire in later adulthood. *Behavior Genetics, 33*, 171-180. doi: 10.1023/a:1022558002760

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., Kerestes, G., Brkovic, I., Butkovic, A., & Denissen, J. (2014). Reciprocal associations between parenting challenges and parents' personality development in young and middle adulthood. *European Journal of Personality, 28,* 168-179. doi: 10.1002/per.1932

Iacono, W. G., & McGue, M. (2002). Minnesota Twin Family Study. *Twin Research, 5,* 482-487.

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.

Jeronimus, B. F., Ormel, J., Aleman, A., Penninx, B. W. J. H., & Riese, H. (2013). Negative and positive life events predict small but long-term change in neuroticism. *Psychological Medicine, 43,* 2403-2415. doi: 10.1017/S0033291713000159

Jeronimus, B. F., Riese, H., Sanderman, R., & Ormel, J. (2014). *Journal of Personality and Social Psychology, 107,* 751-764. doi: 10.1037/a0037009

Jokela, M., Kivimaki, M., Elovainio, M., & Keltikangas-Jarvinen, L. (2009). Personality and having children. A two-way relationship. *Journal of Personality and Social Psychology, 96,* 218-230. doi: 10.1037/a0014058

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.

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

Kandler, C., Riemann, R., Spinath, F. M., & Angleitner, A. (2010). Sources of variance in personality facets: A multiple-rater twin study of self-peer, peer-peer, and self-self (dis)agreement. *Journal of Personality, 78,* 1565-1594.

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

Kern, M. L., & Friedman, H. S. (2008). Do conscientious individuals live longer? A quantitative review. *Health Psychology, 27,* 505-512.

Kim-Cohen, J., Caspi, A., Taylor, A., Williams, B., Newcombe, R., Craig, I. W., & Moffitt, T. E. (2006). MAOA, maltreatment, and gene-environment interaction

predicting children's mental health: New evidence and a meta-analysis. *Molecular Psychiatry, 11,* 903-913. doi: 10.1038/sj.mp.4001851

Lehnart, J., Neyer, F. J., & Eccles, J. (2010). Long-term effects of social investment: The case of partnering in young adulthood. *Journal of Personality, 78,* 639-669. doi: 10.1111/j.1467-6494.2010.00629.x

Leikas, S., & Salmela-Aro, K. (2015). Personality trait changes among young Finns: The role of life events and transitions. *Journal of Personality, 83,* 117-126. doi: 10.1111/jopy.12088

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

Loehlin, J. C., & Martin, N. G. (2001). Age changes in personality traits and their heritabilities during the adult years: evidence from Australian Twin Registry samples. *Personality and Individual Differences, 30*, 1147-1160. doi: 10.1016/s0191-8869(00)00099-4

Lubke, G. H., Hottenga, J. J., Walters, R., Laurin, C., de Geus, E. J., Willemsen, G., . . . Boomsma, D. I. (2012). Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. *Biological Psychiatry, 72,* 707-709. doi: 10.1016/j.biopsych.2012.02.011

Lucas, R. E., & Donnellan, M. B. (2011). Personality development across the life span: Longitudinal analyses with a national sample from Germany. *Journal of Personality and Social Psychology, 101,* 847-861. doi: 10.1037/a0024298

Ludtke, O., Roberts, B. W., Trautwein, U., & Nagy, G. (2011). A random walk down University Avenue: Life paths, life events, and personality trait change at the transition to university life. *Journal of Personality and Social Psychology, 101,* 620-637 doi: 10.1037/a0023743

Luhmann, M., Orth, U., Specht, J., Kandler, C., & Lucas, R. E. (2014). Studying changes in life circumstances and personality: It's about time. *European Journal of Personality, 28,* 256-266. doi: 10.1002/per.1951

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., . . . Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature, 461*, 747-753.

Matteson, L. K., McGue, M., & Iacono, W. G. (2013). Shared environmental influences on personality: A combined twin and adoption approach. *Behavior Genetics, 43,* 491-504. doi: 10.1007/s10519-013-9616-8

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: Theory and Research* (3rd ed.). New York: Guildford.

McCrae, R. R., Costa, P. T., Ostendorf, F., Angleitner, A., Hrebickova, M., Avia, M. D., . . . Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology, 78,* 173-186. doi: 10.1037/0022-3514.78.1.173

McGue, M., Keyes, M., Sharma, A., Elkins, I., Legrand, L., Johnson, W., & Iacono, W. G. (2007). The environments of adopted and non-adopted youth: Evidence on

range restriction from the Sibling Interaction and Behavior Study (SIBS). *Behavior Genetics, 37*, 449-462.

McGue, M., Osler, M., & Christensen, K. (2010). Causal inference and observational research: The utility of twins. *Perspectives on psychological Science, 5,* 546-556.

McGue, M., Zhang, Y., Miller, M. B., Basu, S., Vrieze, S., Hicks, B., . . . Iacono, W. G. (2013). A genome-wide association study of behavioral disinhibition. *Behavior Genetics, 43,* 363-373. doi: 10.1007/s10519-013-9606-x

Middeldorp, C. M., Cath, D. C., Beem, A. L., Willemsen, G., & Boomsma, D. I. (2008). Life events, anxious depression, and personality: A prospective and genetic study. *Psychological Medicine, 38,* 1557-1565. doi: 10.1017/S9933291708002985

Miller, M. B., Basu, S., Cunningham, J., Eskin, E., Malone, S. M., Oetting, W. S., . . . McGue, M. (2012). The Minnesota Center for Twin and Family Research genome-wide association study. *Twin Research in Human Genetics, 15,* 767-774. doi: 10.1017/thg.2012.62

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2002). *Mx: Statistical modeling* (6th edn). Department of Psychiatry: Richmond.

Neyer, F. J., & Lehnart, J. (2007). Relationships matter in personality development: Evidence from an 8-year longitudinal study across young adulthood. *Journal of Personality, 75,* 535-568. doi: 10.1111/j.1467-6494.2007.00448.x

Neyer, F. J., Mund, M., Zimmermann, J., & Wrzus, C. (2014). Personality-relationship transactions revisited. *Journal of Personality, 82,* 539-550. doi: 10.1111/jopy.12063

Parker, P. D., Ludtke, O., Trautwein, U., & Roberts, B. W. (2012). Personality and relationship quality during the transition from high school to early adulthood. *Journal of Personality, 80,* 1061-1089. doi: 10.1111/j.1467-6494.2012.00766.x

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>3.3.co;2-e

Peeters, H., Van Gestel, S., Vlietinck, R., Derom, C., & Derom, R. (1998). Validation of a telephone zygosity questionnaire in twins of known zygosity. *Behavioral Genetics, 28,* 159-163.

Power, R. A., & Pluess, M. (2015). Heritability estimates of the Big Five personality traits based on common genetic variants. *Translational Psychiatry, 5,* 1-4. doi: 10.1038/tp.2015.96

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics, 38*, 904-909.

Purcell, S. (2002). Variance components models for gene-environment interaction in twin analysis. *Twin Research, 5,* 554-571. doi: 10.1375/136905202762342035

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.

Rietveld, C. A., Cesarini, D., Benjamin, D. J., Koellinger, P. D., De Neve, J. E., Tiemeier, H., . . . Bartels, M. (2013). Molecular genetics and subjective well-being. *Proceedings of the National Academy of Sciences, 110,* 9692-9697. doi: 10.1073/pnas.1222171110

Risch, N., Herrell, R., Lehrer, T., Liang, K. Y., Eaves, L., Hoh, J., . . . Merikangas, K. R. (2009). Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: A meta-analysis. *Journal of the American Medical Association, 301,* 2462-2471.

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2001). The kids are alright: Growth and stability in personality development from adolescence to adulthood. *Journal of Personality and Social Psychology, 81,* 670-683. doi: 10.1037/0022-3514.81.4.670

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*, 3-25. doi: 10.1037/0033-2909.126.1.3

Roberts, B. W., 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. doi: 10.1037/0033-2909.132.1.1

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: Theory and research* (3rd ed., pp. 375-398). New York, NY: Gilford.

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

Robins, R. W., Caspi, A., & Moffitt, T. E. (2002). It's not just who you're with, it's who you are: Personality and relationship experiences across multiple relationships. *Journal of Personality, 70,* 925-964.

Sadeh, N., Javdani, S., & Verona, E. (2013). Analysis of monoaminergic genes, childhood abuse, and dimensions of psychopathy. *Journal of Abnormal Psychology, 122,* 167-179. doi: 10.1037/a0029866

Saudino, K. J., Pedersen, N. L., Lichtenstein, P., McClearn, G. E., & Plomin, R. (1997). Can personality explain genetic influences on life events? *Journal of Personality and Social Psychology, 72,* 196-206.

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of Big Five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology, 94,* 718-737. doi: 10.1037/0022-3514.94.4.718

Specht, J., Bleidorn, W., Denissen, J. J. A., Hennecke, M., Hutteman, R., Kandler, C., . . .Zimmermann, J. (2014). What drives adult personality development? A

comparison of theoretical perspectives and empirical evidence. *European Journal of Personality, 28,* 216-230. doi: 10.1002/per.1966

Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology, 101,* 862-882. doi: 10.1037/a0024950

Tellegen, A., & Waller, N. G. (2008). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In G. J. Boyle, G. Matthews, and D. H. Saklofske (Eds.), *Handbook of personality theory and testing: Vol. II. Personality measurement and assessment* (pp. 261-292). Thousand Oaks, CA: Sage.

Terracciano, A., Sanna, S., Uda, M., Deiana, B., Usala, G., Busonero, F., . . . Costa, P. T. (2010). Genome-wide association scan for five major dimensions of personality. *Molecular Psychiatry, 15,* 647-656. doi: 10.1038/mp.2008.113

Trzaskowski, M., Dale, P. S., & Plomin, R. (2013). No genetic influence for childhood behavior problems from DNA analysis. *Journal of the American Academy of Child & Adolescent Psychiatry, 52,* 1048-1056.

Turkheimer, E., & Waldron, M. (2000). Nonshared environment: A theoretical, methodological, and quantitative review. *Psychological Bulletin, 126,* 78-108.

Uher, R. (2013). The changing understanding of the genetic and environmental causes of mental illness. *Canadian Journal of Psychiatry, 58,* 67-68.

Vaidya, J. G., Gray, E. K., Haig, J. R., Mrocek, D. K., & Watson, D. (2008). Differential stability and individual growth trajectories of big five and affective traits during young adulthood. *Journal of Personality, 76,* 267-304. doi: 10.1111/j.1467-6494.2007.00486.x

van den Berg, S. M., de Moor, M. H. M., McGue, M., Pettersson, E., Terracciano, A., Verweij, K. J. H., . . . Boomsma, D. I. (2014). Harmonization of neuroticism and extraversion phenotypes across inventories and cohorts in the Genetics of Personality Consortium: An application of Item Response Theory. *Behavior Genetics, 44,* 295-313. doi: 10.1007/s10519-014-9654-x

Verweij, K. J. H., Yang, J., Lahti, J., Veijola, J., Hintsanen, M., Pulkki-Raback, L., . . . Zietsch, B. P. (2012). Maintenance of genetic variation in human personality: Testing evolutionary models by estimating heritability due to common causal variants and investigating effect of distant inbreeding. *Evolution, 66*, 3238-3251. doi: 10.1111/j.1558-5646.2012.01679.x

Viken, R. J., Rose, R. J., Kaprio, J., & Koskenvuo, M. (1994). A developmental genetic analysis of adult personality: Extroversion 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

Vrieze, S. I., Iacono, W. G., & McGue, M. (2012). Confluence of genes, environment, development, and behavior in a post-GWAS world. *Developmental Psychopathology, 24,* 1195-1214. doi: 10.1017/S0954579412000648

Vukasovic, T., & Bratko, D. (2015). Heritability of personality: A meta-analysis of behavior genetic studies. *Psychological Bulletin, 141,* 769-785.

Wagner, J., Becker, M., Ludtke, O., & Trautwein, U. (2015). The first partnership experience and personality development: A propensity score matching study in young adulthood. *Social Psychological and Personality Science, 6,* 455-463. doi: 10.1177/1948550614566092

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

Yang, J., Manolio, T. A., Pasquale, L. R., Boerwinkle, E., Caporaso, N., Cunningham, J. M., . . . Visscher, P. M. (2011b). Genome-partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics, 43,* 519-525. doi: 10.1038/ng.823

Zheng, H. F., Ladouceur, M., Greenwood, C. M. T., & Richards, J. B. (2012). Effect of genome-wide genotyping and reference panels on rare variants imputation. *Journal of Genetics and Genomics, 39*, 545-550.