Behav Genet
DOI 10.1007/s10519-017-9875-x

ORIGINAL RESEARCH

The Impact of Variation in Twin Relatedness on Estimates of Heritability and Environmental Influences

Chang Liu[1] · **Peter C. M. Molenaar**[2] · **Jenae M. Neiderhiser**[1]

Received: 18 September 2016 / Accepted: 3 October 2017

Abstract By taking advantage of the natural variation in genetic relatedness among identical (monozygotic: MZ) and fraternal (dizygotic: DZ) twins, twin studies are able to estimate genetic and environmental contributions to complex human behaviors. Recently concerns have been raised about the accuracy of twin studies in light of findings of genetic and epigenetic changes in twins. One of the concerns raised is that MZ twins are not 100% genetically and epigenetically similar because they show variations in their genomes and epigenomes leading to inaccurate estimates of heritability. This article presents findings from a simulation study that examined the degree of bias in estimates of heritability and environmentality when the genetic and epigenetic similarity of MZ twins differs from 1.00 and when the genetic and epigenetic similarity of DZ twins differs from 0.50. The findings suggest that in the standard biometric model when MZ or DZ twin similarity differs from 1.00 or 0.50, respectively, the variance that should be attributed to genetic influences is instead attributed to nonshared environmental influences, thus deflating the estimates of genetic influences and inflating the estimates of nonshared environmental influences. Although estimates of genetic and nonshared environmental influences from the standard biometric model were found to deviate from "true" values, the bias was usually smaller than 10% points indicating that the interpretations of findings from previous twin studies are mostly correct.

Keywords Twin studies · Standard biometric model · Genetic and epigenetic similarity · Heritability estimate · Simulation study

Edited by Pete Hatemi.

Electronic supplementary material The online version of this article (doi:10.1007/s10519-017-9875-x) contains supplementary material, which is available to authorized users.

✉ Chang Liu
cxl492@psu.edu

Peter C. M. Molenaar
pxm21@psu.edu

[1] Department of Psychology, The Pennsylvania State University, 431 Moore Building, University Park, PA 16802, USA

[2] Department of Human Development and Family Studies, The Pennsylvania State University, 415 BBH Building, University Park, PA 16802, USA

Introduction

Behavioral genetic designs provide a powerful tool for estimating the relative contributions of heritable and environmental influences on complex human traits (Plomin et al. 2013). The most commonly used family-based behavioral genetic design is the twin study, which compares the similarity of monozygotic (MZ) and dizygotic (DZ) twin pairs. Specifically, twin studies partition the variance of a measured phenotype into additive genetic (A), nonadditive genetic (D), shared environmental (C), and nonshared environmental (E) influences. Findings from twin studies have been shown to be highly consistent with almost every human trait influenced by genetic factors, including physical, medical, psychological and behavioral characteristics (Plomin et al. 2016; Polderman et al. 2015; Turkheimer 2000). In general, twin studies play a critical role in understanding the nature of complex human behaviors.

Recently, concerns have been raised about twin studies and the basic assumptions of twin studies based, in part, on findings from epigenetic and molecular genetic studies (e.g. Burt and Simons 2014; Charney 2012; Gringras

Published online: 02 November 2017

and Chen 2001; Lerner 2004). Epigenetic and molecular genetic studies have found that there is variation in MZ twin genomes and epigenomes, suggesting that they are not 100% identical genetically and epigenetically (e.g. Bruder et al. 2008; Charney 2012). Furthermore, Charney (2012) has described several factors that may result in less genetic and epigenetic similarity in MZ and DZ twins than expected, including genetic mechanisms (e.g., retrotransposons, copy number variations), epigenetic modifications (e.g., histone regulation, DNA methylation), and reported effects of these mechanisms (see Table 1). For example, one study reported large-scale copy-number variations (CNV) among 19 MZ twin pairs with either concordant or discordant phenotypes, indicating genotypic diversity within MZ twin pairs (Bruder et al. 2008). There is also preliminary evidence that genetic differences in MZ twins may explain discordant phenotypes in MZ twins based on studies examining one or two pairs of MZ twins with specific disorders such as Williams syndrome (Castorina et al. 1997) and schizophrenia (Tsujita et al. 1998).

Epigenetic mechanisms regulate the transcriptional activity of genes without changing the DNA sequence (Goldberg et al. 2007; Jaenisch and Bird 2003). Several studies have found epigenetic differences within pairs of MZ twins and cross-sectional and longitudinal studies have indicated that epigenetic discordance of MZ twin pairs increases over time, suggesting inter- and intra-variability among MZ twin pairs (e.g., Fraga et al. 2005; Kaminsky et al. 2009; Ollikainen et al. 2010; Wong et al. 2010). Studies comparing epigenetic discordance rates among MZ and DZ twin pairs have suggested genetic and environmental (including stochastic) influences on epigenetic variation (Bell and Spector 2011; Van Dongen et al. 2016; van; Dongen et al. 2012) and differences in epigenetic profiles have been found to account for MZ twin discordance for a wide range of phenotypes, including schizophrenia and bipolar disorder (Castellani et al. 2015; Dempster et al. 2011; Kuratomi et al. 2008), although some of the findings are from studies examining one or two pairs of MZ twins. Findings like these have raised the question of whether MZ twins can truly be described as genetically and epigenetically identical (Charney 2012; Gringras and Chen 2001). This is a critical concern, as the classical twin design relies upon MZ twins sharing all of their segregating genes and DZ twins sharing half, on average. One of the concerns raised is that the variation in the genetic and epigenetic similarity in MZ twins will result in inaccurate estimates of heritability, because the MZ twins are not, in truth, genetically and epigenetically identical.

Interestingly, there are also published findings that differences among the genome sequences and epigenome of MZ twins are rare (Baranzini et al. 2010; van Dongen et al. 2012; Veenma et al. 2012; Weber-Lehmann et al. 2014). For example, one study examined genome sequence variations among three MZ twin pairs and failed to find evidence for any replicable differences (Baranzini et al. 2010). Although in the same study, epigenetic differences were evident in these three MZ twin pairs, these differences cannot explain disease discordance (Baranzini et al. 2010). Some researchers have argued that genetic and epigenetic differences between MZ twins are not related to heritability because these differences are acquired and not in the inherited DNA (Miller et al. 2012). These advocates continue to argue that MZ twins are not very different in measured genotypes, although they have acknowledged that DNA sequence (e.g., CNVs) and functional differences (e.g., methylation patterns) can be potential sources for MZ twin discordance. In addition, other researchers have conceptualized epigenetics as a source of random effects on the phenotypes independent from standard genetic and environmental influences estimated from the standard biometric model (Bell and Spector 2011; Dolan et al. 2015). In a simulation study, Dolan et al. (2015) examined the consequences of ignoring this random effect and found that ignoring this randomness resulted in small but noticeable influences on the parameter estimates of the standard biometric models.

In sum, there are studies that have found genetic and epigenetic differences between MZ twins, although it is not clear how these differences may affect heritability estimates. Generally, the belief has been that if the actual genetic and epigenetic similarity between MZ twins is less than 100% heritability estimates are likely to be inflated (Handel et al. 2010). The corresponding effects of these epigenetic and genetic differences on DZ twin similarity have not been considered. In the classical twin design, it is assumed that complex human traits are explained by genetic (e.g., additive and nonadditive) and environmental (e.g., shared and nonshared) influences. However, according to the findings from epigenetic and molecular genetic studies reviewed above, not only genetic influences, but also epigenetic (e.g., epigenetic regulation) influences affect human behaviors. In the current study, we assume that epigenetic influences work together with additive genetic influences to affect human behaviors. Therefore, in the current study, A is denoted as both additive genetic and epigenetic influences and thus the similarity of A in MZ twins refers to the additive genetic and epigenetic similarity in MZ twins. It should be noted that environments and stochastic factors also influence epigenetic effects, which is, however, not modeled in current study. The environments and stochastic factors were modeled in Dolan et al. (2015), where epigenetic effects were moderated by both genetic and environmental influences (including stochastic errors). Because heritability and environmentality (shared and nonshared environmental influences) are dependent upon the observed similarity of MZ and DZ twins and are estimated using the assumed genetic and epigenetic similarity, understanding the impact of deviations from the

Table 1 Genetic and epigenetic mechanisms that make MZ and DZ twins to be less similar than expected

Genetic mechanisms	Details	Quantification of reported effects
Retrotransposition	The retrotransposition events during nervous system development can increase nuclear DNA (nDNA) and change DNA sequence (Martin 2009; Singer et al. 2010), inducing intertwin and intratwin genetic heterogeneity of their brains	
Copy number variations (CNVs)	Evidence has shown that the CNVs existed within MZ and DZ twin pairs, leading to intertwine genetic differences (Bruder et al. 2008; Scheet et al. 2012)	E.g., the copy number differences in CNVs between MZ co-twins or DZ co-twins is estimated as 10% per twin pair (Bruder et al. 2008; Czyz et al. 2012; Scheet et al. 2012)
Aneuploidy (e.g., chromosomal aneuploidy)	The neuronal chromosomal aneuploidy is highly stochastic, which contributes to the neuronal chromosomal differences in MZ twins and more than 50% discordance of neuronal chromosomes in DZ twins (Charney 2012)	E.g., the percentage of aneuploidy neural cells is estimated as 10% in the normal adult brain (Iourov et al. 2006; Rehen et al. 2005)
Mitochondrial DNA (mtDNA)	Mitochondrial DNA is stochastically divided with the first mitotic cell division, leading to discordance of mtDNA in MZ twins (Charney 2012; Montier et al. 2009)	
Epigenetic mechanisms/modifications		
[e.g., Histone Regulation, DNA methylation and Non-coding microRNAs (miRNAs)]	Epigenetic mechanisms regulate the transcriptional activity of genes without changing the DNA sequence (Goldberg et al. 2007; Jaenisch and Bird 2003). Several studies have found substantial epigenetic differences within pairs of MZ twins at unique genomic regions, selected tissues samples and in the whole epigenome (Fraga et al. 2005; Heijmans et al. 2007; Kaminsky et al. 2009), which may be associated with MZ twin discordance in a wide range of phenotypes (Castellani et al. 2015; Dempster et al. 2011; Kuratomi et al. 2008)	E.g., the DNA methylation profiles are up to 82% discordant in DZ twin pairs and up to 54% discordant in MZ twin pairs (Ollikainen et al. 2010)

expected genetic and epigenetic similarity in both MZ and DZ twins is important for understanding the impact, if any, of these deviations on these estimates.

The goals of the current study are to (1) examine the impact of genetic and epigenetic differences on both MZ and DZ twin similarities on heritability estimates and (2) discuss the implications for classical twin research. Specifically, we illustrate the degree of bias in estimates of heritability and environmentality when genetic and epigenetic similarity of MZ twins differs from 1.00 as well as when genetic and epigenetic similarity of DZ twins differs from 0.50. To estimate the impact of violating this assumption, we have conducted a series of simulations allowing for variation in the "true" genetic and epigenetic similarity of both types of twins. We have chosen uniform reduction of genetic and epigenetic similarity among MZ twins and DZ twins respectively, because of the simplicity of fitting the biometric model. However, it should be noted that, the current model can also handle data with diversiform reductions of MZ and DZ twin correlations (additional information available on request).

The Simulation Study

One objective of classical twin studies is to partition the variance of a measured phenotype (Y) into three components: (1) additive genetic influences, (2) shared environmental influences, and (3) nonshared environmental influences. One thing to note, in the current study, the variance of a measured phenotypes is partitioned into additive genetic and epigenetic influences, shared environmental and nonshared environmental influences. For the simplicity of the models used in this simulation study, we do not include nonadditive genetic effects (epistasis and dominance) in the model. Let $\mathbf{y}_{ijkm}$ denote the observed phenotypic score at the mth observed variable ($m = 1, 2, \ldots, M$) for the jth member within twin pairs ($j = 1,2$) of the ith twin pair ($i = 1, 2, \ldots, N$) of type k ($k = 1$ for MZ and $k = 2$ for DZ). Then the biometric model is defined as:

$$\mathbf{y}_{ijkm} = \mathbf{a}_m\mathbf{A}_{ijk} + \mathbf{c}_m\mathbf{C}_{ijk} + \mathbf{e}_m\mathbf{E}_{ijk} + \boldsymbol{\varepsilon}_{ijkm} \quad (1)$$

In the equation, $\mathbf{A}_{ijk}$ is the additive genetic and epigenetic factor score of the jth member of the ith twin pair of type k; $\mathbf{C}_{ijk}$ is the shared environmental factor score of the jth member of the ith twin pair of type k and $\mathbf{E}_{ijk}$ is the nonshared environmental factor score of the jth member of the ith twin pair of type k. $\mathbf{a}_m$, $\mathbf{c}_m$ and $\mathbf{e}_m$ denote the factor loadings of the mth phenotype on, respectively, the additive genetic and epigenetic factor $\mathbf{A}$, the shared environmental factor $\mathbf{C}$ and specific (nonshared) environmental factor $\mathbf{E}$, while $\boldsymbol{\varepsilon}_{ijkm}$ denotes measurement error.

In the standard biometric model, the genetic and epigenetic factor correlation (genetic and epigenetic similarity) across MZ twin pairs is 1.00 and across DZ twin pairs is 0.50: $\text{cor}(\mathbf{A}_{i1k}, \mathbf{A}_{i2k}) = 1.00$ if $k = 1$ and $\text{cor}(\mathbf{A}_{i1k}, \mathbf{A}_{i2k}) = 0.50$ if $k = 2$. The correlation of shared environmental factors is 1.00 and the nonshared environmental factors are not correlated for both MZ twin pairs and DZ twin pairs: $\text{cor}(\mathbf{C}_{i1k}, \mathbf{C}_{i2k}) = 1.00$ for $k = 1$ and $k = 2$ and $\text{cor}(\mathbf{E}_{i1k}, \mathbf{E}_{i2k}) = 0$ for $k = 1$ and $k = 2$. Figure 1 illustrates the path diagram for a standard biometric model. As shown in Fig. 1, the correlation linking the two A latent factors is set to 1.00 for MZ twins and 0.50 for DZ twins, which defines the expected variance–covariance matrix for MZ and DZ twins. The correlation linking the two C latent factors is set to 1.00 for both MZ and DZ twins when they are reared in the same home. There is no correlation between the two E latent factors

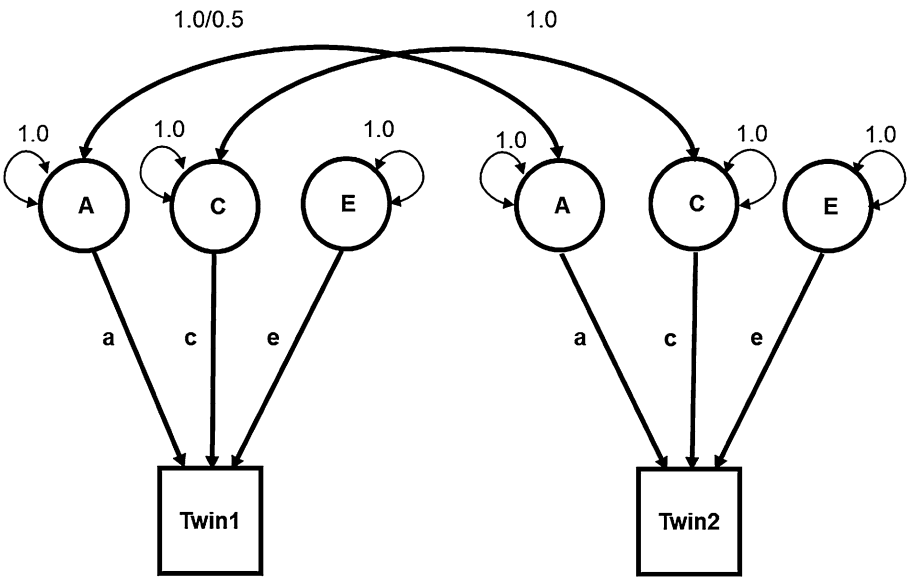

Fig. 1 ACE path diagram. This path diagram is equivalent to the equation of the standard biometric model. Factor loadings (or path coefficients: *a, c*, and *e*) rather than variance components (*A, C*, and *E*) are estimated

because, by definition, nonshared environmental influences account for differences between members of twin pairs.

In contrast to the fixed genetic and epigenetic factor correlations in the standard biometric model, the alternative model sets the genetic and epigenetic factor correlations of MZ and DZ twins to be free parameters that need to be estimated. The correlations of shared and nonshared environmental factors of MZ and DZ twins are set to be the same as correlations in the standard biometric model. When fitting multivariate twin model, the alternative model is an identifiable model, which generates the standard deviation for each parameter estimate.

In the current simulation study, our aim is to examine the degree of bias in estimates of genetic, shared environmental and nonshared environmental influences when genetic and epigenetic factor correlation of MZ twins differs from 1.00 and when genetic and epigenetic factor correlation of DZ twins differs from 0.50. Thus, in generating the data, we set the "true" level of genetic and epigenetic factor correlation of MZ twins to vary between 1.00 and 0.80 (e.g., 1.00, 0.90 and 0.80) and the "true" level of genetic and epigenetic factor correlation of DZ twins to vary between 0.50 and 0.40 (e.g., 0.50, 0.45 and 0.40): $\text{cor}(\mathbf{A}_{i1k}, \mathbf{A}_{i2k}) = 1.00$, 0.90 or 0.80 if $k = 1$ and $\text{cor}(\mathbf{A}_{i1k}, \mathbf{A}_{i2k}) = 0.50$, 0.45 or 0.40 if $k = 2$. We choose these values based on the reported genetic and epigenetic discordance within MZ and DZ twin pairs (see Table 1), which is moderate compared to some of the reported effects.

We include five phenotypes (y) in the model, thus $m = 1$, 2, 3, 4, 5, which is a multivariate twin model with five phenotypes. In generating the data based on the model described above, we assign numerical values to the fixed parameters in the model. The variances of A, C and E are fixed at 1.00 and the variance of the measurement errors are set equal to 0.70: $\text{var}[\boldsymbol{\varepsilon}_{ijk1}] = \text{var}[\boldsymbol{\varepsilon}_{ijk2}] = \text{var}[\boldsymbol{\varepsilon}_{ijk3}] = \text{var}[\boldsymbol{\varepsilon}_{ijk4}] = \text{var}[\boldsymbol{\varepsilon}_{ijk5}] = 0.70$. The factor loadings of the genetic and epigenetic factors, shared environmental factors and non-shared environmental factors are chosen as follows:

	a	c	e
Phenotype 1	2.0	1.0	1.0
Phenotype 2	2.0	1.5	1.0
Phenotype 3	2.0	1.2	1.8
Phenotype 4	1.0	1.5	1.4
Phenotype 5	0.8	1.0	2.0

These factor loadings are unstandardized and can represent factor loadings for A, C and E reported from various twin studies. The reliability for each phenotype is 0.90, 0.91, 0.93, 0.88 and 0.89 respectively. Last, there are nine combinations (conditions) of "true" genetic and epigenetic factor correlations across MZ twin pairs and across DZ twin pairs, where both factors (different genetic and epigenetic MZ correlations and different genetic and epigenetic DZ correlations) are fully crossed. 900 data sets are generated—each with 900 MZ twin pairs and 900 DZ twin pairs—for each condition.

The next step is to analyze the generated data using both the standard biometric model with fixed genetic and epigenetic factor correlations (1.00 for MZ twins and 0.50 for DZ twins) and the alternative model with freely estimated genetic and epigenetic factor correlations. The heritability, shared environmental influence and nonshared environmental influence are estimated for five phenotypes under each combination of "true" genetic and epigenetic factor correlations across MZ twin pairs and across DZ twin pairs described above using both models. The standardized factor estimates are squared to provide the heritability (a^2), shared environmental (c^2), and nonshared environmental (e^2) estimates. The NONLIS source code used in fitting the standard biometric model and the alternative model is specified in the appendix. For more information about the NONLIS program, see https://quantdev.ssri.psu.edu/. Last, the estimated heritability, shared environmental influence and nonshared environmental influence from the standard biometric model and from the alternative model are then compared with the "true" heritability, shared environmental influence and nonshared environmental influence used in generating the data. To examine whether the estimated heritability, shared environmental influence and nonshared environmental influence under each condition are significantly different from "true" values, we also report the standard deviation of each estimate.

Results

Results From the Standard Biometric Model

We report the degree of bias in parameter estimates from the standard biometric model when compared with the "true" values and the standard deviation of each parameter estimate. We also examine the effects of various factors on the degree of bias in parameter estimates. The "true" values and estimates of genetic, shared environmental and nonshared environmental factors from the standard biometric model for each phenotype are specified in the appendix.

Effects of Different Genetic and Epigenetic Factor Correlations of MZ Twins

For phenotype 1, as shown in Fig. 2, when genetic and epigenetic factor correlation between MZ twins is 1.00 and between DZ twins is 0.50, the a^2, c^2 and e^2 estimates do not differ from the "true" values. However, under other

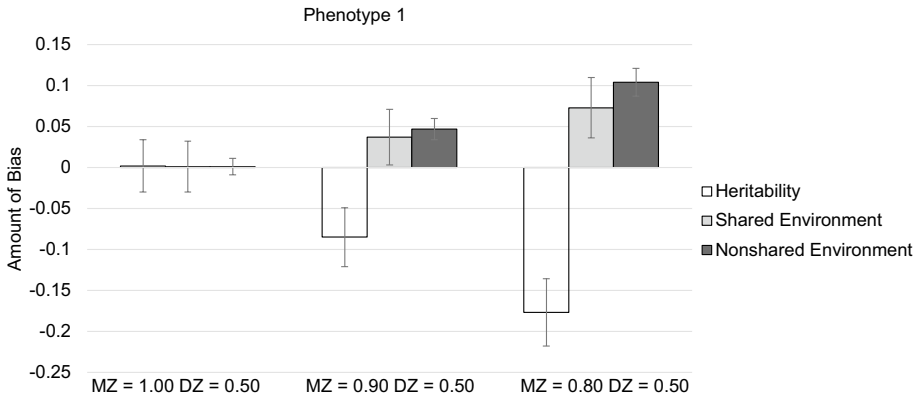

Fig. 2 Bar graph showing the amount of bias in heritability, shared environment, and nonshared environment estimates from the standard biometric model for phenotype 1 when compared with the "true" values, when genetic and epigenetic factor correlation between MZ twins is 1.00, 0.90 and 0.80. Standard deviations of parameter estimates from the standard biometric model are presented in the figure by the error bars attached to each column

conditions where genetic and epigenetic factor correlation between MZ twins is smaller than 1.00, all the estimates of a^2 and e^2 are significantly different from "true" a^2 and e^2 values. Specifically, a^2 is underestimated compared to the "true" a^2 value. The degree of bias in the a^2 estimate increases as the genetic and epigenetic factor correlation of MZ twins decreases. In contrast to a^2, e^2 is overestimated, the degree of bias in which increases as the genetic and epigenetic factor correlation of MZ twins decreases. Similar to e^2, c^2 is overestimated, compared to the "true" value for c^2, although the bias is significant only when genetic and epigenetic similarity between MZ twins is 0.80, but not 0.90. For example, a 0.1 decrease in the genetic and epigenetic correlation of MZ twins leads to roughly 8% points decrease in the heritability estimate, 5% points increase in the nonshared environmental influence estimate and 3% points increase in the shared environmental influence estimate.

Effects of Different Genetic and Epigenetic Factor Correlations of DZ Twins

Take phenotype 2 as an example. As shown in Fig. 3, when genetic and epigenetic factor correlation between MZ twins is 1.00 and between DZ twins is 0.50, 0.45 or 0.40, the a^2, c^2 and e^2 estimates are within 95% confidence interval about the "true" a^2, c^2 and e^2 values. When genetic and epigenetic factor correlation between DZ twins is 0.45 or 0.40, although there is a trend that a^2 is overestimated and c^2 is underestimated compared to the "true" values, the differences between the estimates and "true" values are not significant.

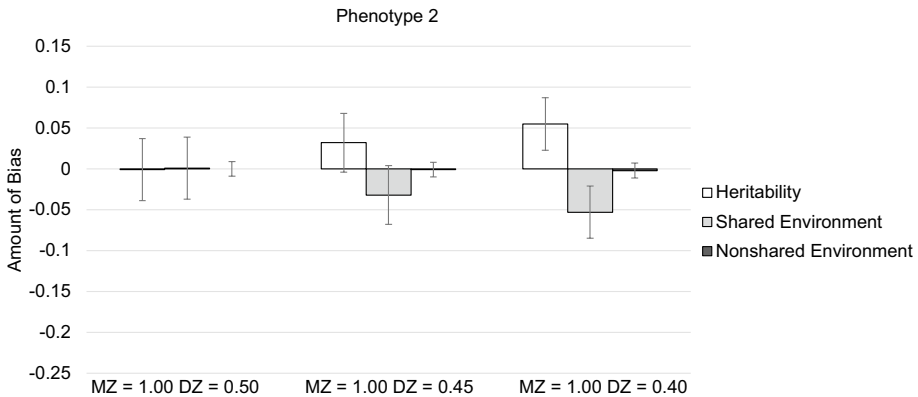

Fig. 3 Bar graph showing the amount of bias in heritability, shared environment, and nonshared environment estimates from the standard biometric model for phenotype 2 when compared with the "true" values, when genetic and epigenetic factor correlation between DZ twins is 0.50, 0.45 and 0.40. Standard deviations of parameter estimates from the standard biometric model are presented in the figure by the error bars attached to each column

Effects of Different Combinations of Genetic and Epigenetic Factor Correlations of MZ and DZ Twins

As an example, we present the bias in parameter estimates for phenotypes 1 and 2 in Figs. 4 and 5. When genetic and epigenetic factor correlations of MZ and DZ twins are 0.90 and 0.45/0.40 respectively, a^2 and c^2 estimates are not biased, while e^2 estimate is significantly larger than "true" value. As discussed before, a 0.1 decrease in MZ twins' genetic and epigenetic similarity leads to the underestimate of a^2. In contrast, a 0.1 decrease in DZ twins' genetic and epigenetic similarity leads to the trend in the overestimate of a^2. Thus, the bias in a^2 estimate is, in fact, smaller with 0.1 decrease in the genetic and epigenetic similarities of both MZ and DZ twins than with 0.1 decrease in the genetic and epigenetic similarity of only MZ twins, as indicated in the current finding with nonsignificant deviation of a^2 estimate from "true" value. However, when genetic and epigenetic factor correlation between MZ twins decreases to 0.80, a^2 is underestimated and e^2 is overestimated. c^2 estimate is somewhat inflated, although the bias is not significant.

Effects of Different Magnitudes of Genetic and Nonshared Environmental Influences

In order to illustrate the effects of different magnitudes of genetic and nonshared environmental factors on the degree of bias in parameter estimates, we have compared the degree of bias in the estimates of a^2, c^2 and e^2 for phenotypes 1 and 5 (see Table 2). For phenotype 1, the "true" values for a^2,

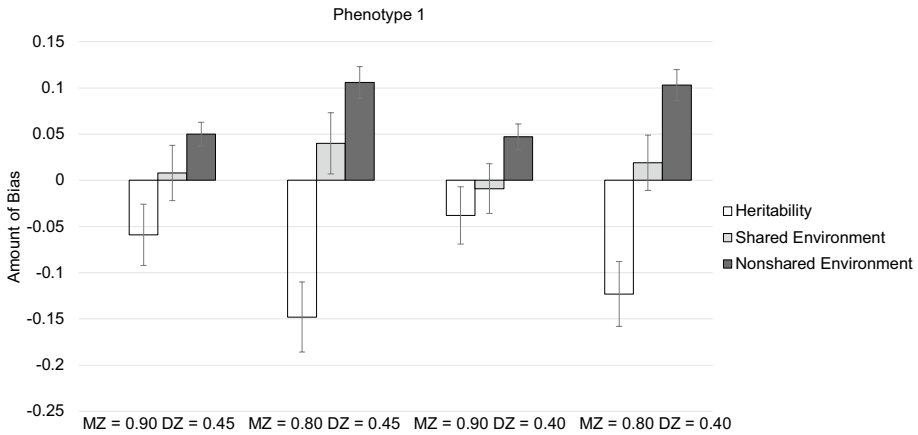

Fig. 4 Bar graph showing the amount of bias in heritability, shared environment, and nonshared environment estimates from the standard biometric model for phenotype 1 when compared with the "true" values, when genetic and epigenetic factor correlation between MZ twins is 0.9 or 0.8 and that between DZ twins is 0.45 or 0.40. Standard deviations of parameter estimates from the standard biometric model are presented in the figure by the error bars attached to each column

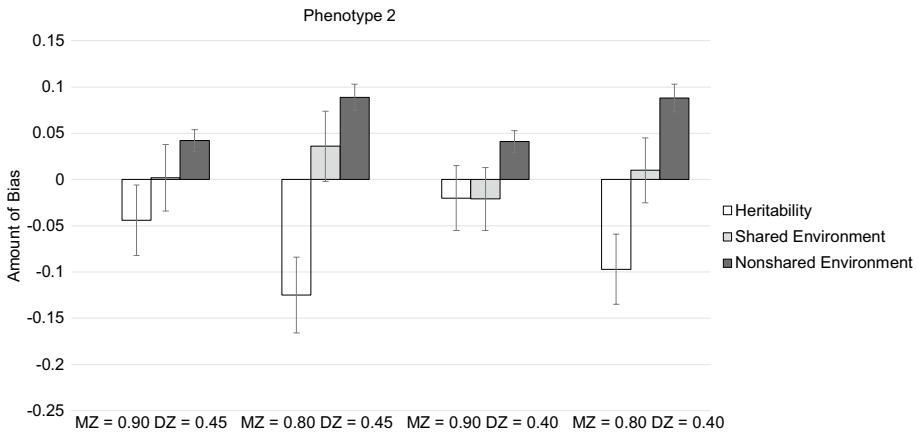

Fig. 5 Bar graph showing the amount of bias in heritability, shared environment, and nonshared environment estimates from the standard biometric model for phenotype 2 when compared with the "true" values, when genetic and epigenetic factor correlation between MZ twins is 0.9 or 0.8 and that between DZ twins is 0.45 or 0.40. Standard deviations of parameter estimates from the standard biometric model are presented in the figure by the error bars attached to each column

Table 2 The amount of bias in heritability, shared environment, and nonshared environment estimates from the standard biometric model for phenotypes 1 and 5 when compared with the "true" values

Amount of bias	Phenotype 1			Phenotype 5		
Genetic and epigenetic factor correlation	a^2 (std.)	c^2 (std.)	e^2 (std.)	a^2 (std.)	c^2 (std.)	e^2 (std.)
MZ = 0.90 DZ = 0.50	−0.085 (0.036)	0.037 (0.034)	0.047 (0.013)	−0.037 (0.017)	0.026 (0.022)	0.012 (0.019)
MZ = 0.80 DZ = 0.50	−0.177 (0.041)	0.073 (0.037)	0.104 (0.017)	−0.068 (0.015)	0.047 (0.023)	0.022 (0.019)
MZ = 1.00 DZ = 0.45	−0.024 (0.030)	0.024 (0.028)	−0.001 (0.010)	0.012 (0.018)	−0.010 (0.021)	−0.001 (0.018)
MZ = 0.90 DZ = 0.45	−0.059 (0.033)	0.008 (0.030)	0.050 (0.013)	−0.025 (0.017)	0.015 (0.021)	0.012 (0.019)
MZ = 0.80 DZ = 0.45	−0.148 (0.038)	0.040 (0.033)	0.106 (0.017)	−0.060 (0.015)	0.038 (0.023)	0.023 (0.019)
MZ = 1.00 DZ = 0.40	0.042 (0.026)	−0.041 (0.025)	−0.003 (0.010)	0.022 (0.018)	−0.017 (0.020)	−0.004 (0.018)
MZ = 0.90 DZ = 0.40	−0.038 (0.031)	−0.009 (0.027)	0.047 (0.014)	−0.015 (0.017)	0.007 (0.021)	0.010 (0.019)
MZ = 0.80 DZ = 0.40	−0.123 (0.035)	0.019 (0.030)	0.103 (0.017)	−0.052 (0.016)	0.030 (0.022)	0.022 (0.020)

Std. standard deviation of heritability, shared environment and nonshared environment estimates, a^2 additive genetic variance component, c^2 shared environmental variance component, e^2 nonshared environmental variance component

c^2 and e^2 are 0.667, 0.167 and 0.167, respectively. For phenotype 5, the "true" values for a^2, c^2 and e^2 are 0.113, 0.177 and 0.709. Hence, shared environmental influence accounts for roughly the same amount of the total variance in these two phenotypes. Genetic influence accounts for more of the total variance for phenotype 1 than for phenotype 5, whereas nonshared environmental influence accounts for more of the total variance for phenotype 5 compared to phenotype 1. The comparison of the estimates for phenotypes 1 and 5 indicates that the degree of bias in the estimates of a^2 and e^2 is much larger for phenotype 1 than for phenotype 5 under all the conditions. For example, for phenotype 1, when genetic and epigenetic factor correlation of MZ twins decreases 0.1, the estimate of a^2 reduces roughly 8% points in the standard biometric model. However, for phenotype 5, the same amount of decrease in the genetic and epigenetic factor correlation of MZ twins only leads to a reduction of 3% points in the estimate of a^2 from the standard biometric model, indicating that the magnitude of genetic influence has a larger impact than nonshared environmental influence on the degree of bias in parameter estimates. The findings indicate that the degree of bias in the estimates of a^2 and e^2 is smaller for phenotypes where genetic influence accounts for less of the total variance compared to phenotypes where genetic influence accounts for more of the total variance when the degree of genetic and epigenetic similarity between MZ twins is less than 100%.

Effects of Different Magnitudes of Genetic and Shared Environmental Influences on Phenotypes

For phenotype 3, the "true" values for a^2, c^2 and e^2 are 0.461, 0.166 and 0.373, respectively. For phenotype 4, the "true" values for a^2, c^2 and e^2 are 0.192, 0.432 and 0.376. In comparison of phenotypes 3 and 4 (see Table 3), nonshared environmental influence explains roughly the same amount of the total variance, but genetic influence explains more of the total variance and shared environmental influence explains less of the total variance for phenotype 3 than for phenotype 4. After comparing the degree of bias in the estimates of a^2, c^2 and e^2 for phenotypes 3 and 4, we have found that the deviations of a^2 and e^2 (but not c^2) estimates from the "true" values are larger for phenotype 3 than for phenotype 4 under the condition that genetic and epigenetic similarity between MZ twins is less than 100%. The comparison of phenotype 3 and phenotype 4 suggests that the magnitude of shared environmental influence has smaller impact on the bias of a^2 and e^2 estimates than genetic influence when the degree of genetic and epigenetic similarity between MZ twins is less than 100%. We have also examined phenotypes with high shared environmental influence (0.8), and the results are comparable to those phenotypes with moderate shared environmental influence presented here (additional results available on request).

Results From the Alternative Model

In the alternative model, all the estimated MZ and DZ twins' genetic and epigenetic factor correlations are within 95% confidence intervals about the "true" genetic and epigenetic factor correlations used in generating the data. In addition, under all the conditions (different combinations of "true" genetic and epigenetic factor correlations of MZ and DZ twins), the estimates of genetic, shared environmental and nonshared environmental influences are not different from "true" values for all five phenotypes. Estimates of genetic, shared environmental and nonshared environmental factors from the alternative model for each phenotype are specified in the appendix in Supplementary material 1.

Table 3 The amount of bias in heritability, shared environment, and nonshared environment estimates from the standard biometric model for phenotypes 3 and 4 when compared with the "true" values

Amount of bias	Phenotype 3			Phenotype 4		
Genetic and epigenetic factor correlation	a^2 (std.)	c^2 (std.)	e^2 (std.)	a^2 (std.)	c^2 (std.)	e^2 (std.)
MZ = 0.90 DZ = 0.50	−0.081 (0.031)	0.022 (0.029)	0.059 (0.018)	−0.059 (0.025)	0.028 (0.029)	0.031 (0.018)
MZ = 0.80 DZ = 0.50	−0.165 (0.033)	0.048 (0.030)	0.118 (0.019)	−0.111 (0.023)	0.054 (0.028)	0.058 (0.018)
MZ = 1.00 DZ = 0.45	0.021 (0.028)	−0.021 (0.026)	−0.001 (0.015)	0.024 (0.028)	−0.024 (0.031)	0.000 (0.017)
MZ = 0.90 DZ = 0.45	−0.060 (0.029)	0.001 (0.026)	0.059 (0.017)	−0.037 (0.026)	0.007 (0.030)	0.030 (0.018)
MZ = 0.80 DZ = 0.45	−0.141 (0.032)	0.024 (0.028)	0.117 (0.019)	−0.089 (0.024)	0.033 (0.029)	0.056 (0.019)
MZ = 1.00 DZ = 0.40	0.038 (0.025)	−0.034 (0.022)	−0.003 (0.017)	0.042 (0.026)	−0.039 (0.029)	−0.002 (0.017)
MZ = 0.90 DZ = 0.40	−0.044 (0.028)	−0.014 (0.024)	0.058 (0.019)	−0.023 (0.026)	−0.007 (0.028)	0.030 (0.018)
MZ = 0.80 DZ = 0.40	−0.124 (0.031)	0.007 (0.026)	0.117 (0.020)	−0.076 (0.024)	0.019 (0.028)	0.057 (0.018)

Std. standard deviation of heritability, shared environment and nonshared environment estimates, a^2 additive genetic variance component, c^2 shared environmental variance component, e^2 nonshared environmental variance component

Discussion

In the current simulation study, we examined the degree of bias in parameter estimates from the standard biometric model when the genetic and epigenetic similarity of MZ twins differs from 1.00 as well as when the genetic and epigenetic similarity of DZ twins differs from 0.50. The findings suggest that the variance that should be attributed to genetic influences is instead attributed to nonshared environmental influences. In other words, the bias in the standard model, as compared to a model using "true" genetic and epigenetic similarity of MZ and DZ twins, results in deflated estimates of genetic influences and inflated estimates of nonshared environmental influences. The implications of these findings for classical twin research are discussed below.

First, for phenotypes with high heritability, genetic influences are underestimated (usually no more than 10% points) if the degree of genetic and epigenetic similarity between MZ twins is less than 1.00. In other words, heritability estimates using a standard biometric model for phenotypes with high heritability provide conservative estimates of heritability. The "true" heritability for these phenotypes is likely to be higher than reported. In contrast, violation of this assumption has little impact on phenotypes with low heritability. Estimates for genetic, shared, and nonshared environmental influences on phenotypes with low heritability do not change if the assumption is violated. Accordingly, the interpretations of heritability estimates using the standard biometric model reported in the literature are mostly correct.

Second, nonshared environmental influences are overestimated when MZ twin pairs are not genetically and epigenetically identical. This may help to explain the difficulty in finding systematic sources of nonshared environmental influences despite the fact that twin studies have consistently reported substantial nonshared environmental influences for almost every human trait (Plomin and Daniels 1987). Several genetically informative longitudinal studies, such as the Nonshared Environmental and Adolescent Development (NEAD) project (Reiss et al. 2000), were designed to examine nonshared environmental influences. However, these studies have made little progress in identifying the systematic causes of the nonshared environmental influences (Reiss et al. 2000; Turkheimer and Waldron 2000). One possible explanation is that the causal effect of a single nonshared environmental event is too small to detect; only the cumulative effects of multiple nonshared environmental events can cause noticeable differences between MZ twins (Plomin and Daniels 1987; Turkheimer and Waldron 2000). Another explanation indicated by the current findings is that nonshared environmental influences are, in fact, smaller than suggested by twin research using the standard biometrical model. It should be noted, however, that the bias in estimates of nonshared environmental influences is relatively small, making it less likely that this is the best explanation for the difficulty in identifying systematic sources of nonshared environmental influences.

Third, there is some question about the small effects of shared environmental influences, compared to genetic and nonshared environmental influences, on complex human behaviors reported from previous twin studies. Some researchers have argued that shared environmental influences are underestimated in the biometric models, especially when basic assumptions underlying the biometric models are violated (Burt and Simons 2014; Dickens and Flynn 2001). This has led some to call for abandoning behavioral genetic studies (Burt and Simons 2014). However, studies reporting small effects of shared environmental influences are most often focused on personality and cognitive ability in adults. In contrast, recent studies, including meta-analysis, have examined shared

environmental influences in child and adolescent psychopathology and parenting and found moderate and significant contributions of shared environmental influences (Burt 2009; Klahr and Burt 2014; Neiderhiser et al. 2004). In addition to these studies, the findings from the current simulation analyses do not support the arguments about the underestimation of shared environmental influences. Instead, the current findings indicate that a decrease in the genetic and epigenetic similarity across MZ twins leads to a somewhat overestimation of shared environmental influences. When genetic and epigenetic similarity of DZ twins is smaller than 0.50, although there is a trend that shared environmental influences are underestimated, the bias is small. The findings indicate that the decrease in genetic and epigenetic similarity of DZ twins leads to minor changes in parameter estimates, especially when it is combined with the decrease in genetic and epigenetic similarity of MZ twins.

Fourth, although the bias in parameter estimates from the standard biometric model is small, it is possible to more accurately estimate parameters. In the current study, the alternative model estimated genetic and epigenetic similarities of MZ and DZ twins and applied the estimated genetic and epigenetic similarities to parameter estimates. It was shown in Molenaar et al. (2012) that the genetic and epigenetic similarities are identifiable parameters in the multivariate biometric model. As a result, the alternative model works very well. The genetic and epigenetic similarities of MZ and DZ twins and parameter estimates for each phenotype are nearly the same as the "true" values used in generating the data, suggesting the alternative model with estimated genetic and epigenetic similarities is a promising alternative biometric model.

In sum, although parameter estimates from the standard biometric model deviate from "true" values when the genetic and epigenetic similarity of MZ twins differs from 1.00 and when the genetic and epigenetic similarity of DZ twins differs from 0.50, the bias is not large and the interpretations drawn from the standard biometric model are mostly correct. Thus, the violation of the assumption that MZ twins are genetically identical and DZ twins are 50% genetically identical on average does not invalidate the standard biometric model and heritability estimates.

Compliance with Ethical Standards

Conflict of interest Chang Liu, Peter C. M. Molenaar and Jenae M. Neiderhiser declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human participants or animals performed by any of the authors.

References

Baranzini SE, Mudge J, van Velkinburgh JC, Khankhanian P, Khrebtukova I, Miller NA, Kim RW (2010) Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. Nature 464(7293):1351–1356. doi: 10.1038/nature08990

Bell JT, Spector TD (2011) A twin approach to unraveling epigenetics. Trends Genet. doi: 10.1016/j.tig.2010.12.005

Bruder CE, Piotrowski A, Gijsbers AA, Andersson R, Erickson S, de Ståhl TD, … Poplawski A (2008) Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. Am J Hum Genet 82(3):763–771. doi: 10.1016/j.ajhg.2007.12.011

Burt SA (2009) Rethinking environmental contributions to child and adolescent psychopathology: a meta-analysis of shared environmental influences. Psychol Bull 135(4):608–637. doi: 10.1037/a0015702

Burt CH, Simons RL (2014) Pulling back the curtain on heritability studies: Biosocial criminology in the postgenomic era. Criminology 52(2):223–262. doi: 10.1111/1745-9125.12036

Castellani CA, Laufer BI, Melka MG, Diehl EJ, O'Reilly RL, Singh SM (2015) DNA methylation differences in monozygotic twin pairs discordant for schizophrenia identifies psychosis related genes and networks. BMC Med Genom 8(1):1–12. doi: 10.1186/s12920-015-0093-1

Castorina, P., Selicorni, A., Bedeschi, F., Dalprà, L., & Larizza, L. (1997). Genotype-phenotype correlation in two sets of monozygotic twins with Williams syndrome. Am J Med Genet 69(1):107–111. doi: 10.1002/(SICI)1096-8628(19970303)69:1<107::AID-AJMG21>3.0.CO;2-S

Charney E (2012) Behavior genetics and postgenomics. Behav Brain Sci 35(05):331–358. doi: 10.1017/S0140525X11002226

Czyz W, Morahan JM, Ebers GC, Ramagopalan SV (2012) Genetic, environmental and stochastic factors in monozygotic twin discordance with a focus on epigenetic differences. BMC Med 10(1):93. doi:10.1186/1741-7015-10-93

Dempster EL, Pidsley R, Schalkwyk LC, Owens S, Georgiades A, Kane F, Toulopoulou T (2011) Disease-associated epigenetic changes in monozygotic twins discordant for schizophrenia and bipolar disorder. Hum Mol Genet 20(24):4786–4796. doi: 10.1093/hmg/ddr416

Dickens WT, Flynn JR (2001) Heritability estimates versus large environmental effects: the IQ paradox resolved. Psychol Rev 108(2):346–369. doi: 10.1037/0033-295X.108.2.346

Dolan C, Nivard M, van Dongen J, van der Sluis S, Boomsma D (2015) Methylation as an epigenetic source of random genetic effects in the classical twin design. Adv Genom Genet 5:305–315. doi: 10.2147/AGG.S46909

Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, … Benitez J (2005). Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA 102(30):10604–10609. doi: 10.1073/pnas.0500398102

Goldberg AD, Allis CD, Bernstein E (2007) Epigenetics: a landscape takes shape. Cell 128(4):635–638. doi: 10.1016/j.cell.2007.02.006

Gringras P, Chen W (2001) Mechanisms for differences in monozygous twins. Early Hum Dev 64(2):105–117. doi: 10.1016/S0378-3782(01)00171-2

Handel AE, Ebers GC, Ramagopalan SV (2010) Epigenetics: molecular mechanisms and implications for disease. Trends Mol Med 16(1):7–16. doi: 10.1016/j.molmed.2009.11.003

Heijmans BT, Kremer D, Tobi EW, Boomsma DI, Slagboom PE (2007) Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. Hum Mol Genet 16(5):547–554. doi: 10.1093/hmg/ddm010

Iourov IY, Vorsanova SG, Yurov YB (2006) Chromosomal variation in mammalian neuronal cells: known facts and attractive hypotheses. Int Rev Cytol 249:143–191. doi: 10.1016/S0074-7696(06)49003-3

Jaenisch R, Bird A (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 33:245–254. doi:10.1038/ng1089

Kaminsky ZA, Tang T, Wang S-C, Ptak C, Oh GH, Wong AH, … Petronis A (2009) DNA methylation profiles in monozygotic and dizygotic twins. Nat Genet 41(2):240–245. doi: 10.1038/ng.286

Klahr AM, Burt SA (2014) Elucidating the etiology of individual differences in parenting: a meta-analysis of behavioral genetic research. Psychol Bull 140(2):544–586. doi: 10.1037/a0034205

Kuratomi G, Iwamoto K, Bundo M, Kusumi I, Kato N, Iwata N, Kato T (2008) Aberrant DNA methylation associated with bipolar disorder identified from discordant monozygotic twins. Mol Psychiatry 13(4):429–441. doi: 10.1038/sj.mp.4002001

Lerner RM (2004) Genes and the promotion of positive human development: Hereditarian versus developmental systems perspectives. In: Coll CG, Bearer E, Lerner RM (eds) Nature and nurture: The complex interplay of genetic and environmental influences on human behavior and development. Erlbaum, Mahwah, NJ, p. 1–33

Martin SL (2009) Developmental biology: jumping-gene roulette. Nature 460(7259):1087–1088. doi:10.1038/4601087a

Miller MB, DeYoung CG, McGue M (2012) Assumptions in studies of heritability and genotype–phenotype association. Behav Brain Sci 35(05):372–373. doi: 10.1017/S0140525X12001380

Molenaar PC, Smit DJ, Boomsma DI, Nesselroade JR (2012) Estimation of subject-specific heritabilities from intra-individual variation: iFACE. Twin Res Hum Genet 15(03):393–400. doi: 10.1017/thg.2012.9

Montier L. L. C., Deng JJ, Bai Y (2009) Number matters: control of mammalian mitochondrial DNA copy number. J Genet Genom 36(3):125–131. doi: 10.1016/S1673-8527(08)60099-5

Neiderhiser JM, Reiss D, Pedersen NL, Lichtenstein P, Spotts EL, Hansson K, … Elthammer O (2004) Genetic and environmental influences on mothering of adolescents: a comparison of two samples. Dev Psychol 40(3):335–351. doi: 10.1037/0012-1649.40.3.335

Ollikainen M, Smith KR, Joo E. J.-H., Ng HK, Andronikos R, Novakovic B, … Craig JM (2010) DNA methylation analysis of multiple tissues from newborn twins reveals both genetic and intrauterine components to variation in the human neonatal epigenome. Hum Mol Genet 19(21):4176–4188. doi: 10.1093/hmg/ddq336

Plomin R, Daniels D (1987) Why are children in the same family so different from one another? Behav Brain Sci 10(01):1–16. doi: 10.1017/S0140525X00055941

Plomin R, DeFries JC, Knopik VS, Neiderhiser J (2013) Behavioral genetics. Worth Publishers, New York, NY

Plomin R, DeFries JC, Knopik VS, Neiderhiser JM (2016) Top 10 Replicated Findings From Behavioral Genetics. Perspect Psychol Sci 11(1):3–23. doi: 10.1177/1745691615617439

Polderman TJ, Benyamin B, De Leeuw CA, Sullivan PF, Van Bochoven A, Visscher PM, Posthuma D (2015) Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nat Genet 47(7):702–709. doi: 10.1038/ng.3285

Rehen SK, Yung YC, McCreight MP, Kaushal D, Yang AH, Almeida BS, … Anliker B (2005) Constitutional aneuploidy in the normal human brain. J Neurosci 25(9):2176–2180. doi: 10.1523/JNEUROSCI.4560-04.2005

Reiss D, Neiderhiser JM, Hetherington EM, Plomin R (2000) The relationship code: deciphering genetic and social influences on adolescent development, vol 1. Harvard University Press, Cambridge, MA

Scheet P, Ehli EA, Xiao X, van Beijsterveldt CE, Abdellaoui A, Althoff RR, … Huizenga PE (2012) Twins, tissue, and time: an assessment of SNPs and CNVs. Twin Res Hum Genet 15(06):737–745. doi:10.1017/thg.2012.61

Singer T, McConnell MJ, Marchetto MC, Coufal NG, Gage FH (2010) LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? Trends Neurosci 33(8):345–354. doi: 10.1016/j.tins.2010.04.001

Tsujita T, Niikawa N, Yamashita H, Imamura A, Hamada A, Nakane Y, Okazaki Y (1998) Genomic discordance between monozygotic twins discordant for schizophrenia. Am J Psychiatry 155:422–424. doi: 10.1176/ajp.155.3.422

Turkheimer E (2000) Three laws of behavior genetics and what they mean. Curr Dir Psychol Sci 9(5):160–164. doi: 10.1111/1467-8721.00084

Turkheimer E, Waldron M (2000) Nonshared environment: a theoretical, methodological, and quantitative review. Psychol Bull 126(1):78–108. doi: 10.1037/0033-2909.126.1.78

Van Dongen J, Slagboom PE, Draisma HH, Martin NG, Boomsma DI (2012) The continuing value of twin studies in the omics era. Nat Rev Genet 13(9):640–653. doi: 10.1038/nrg3243

Van Dongen J, Nivard MG, Willemsen G, Hottenga J-J, Helmer Q, Dolan CV, … Breeze CE (2016) Genetic and environmental influences interact with age and sex in shaping the human methylome. Nature communications 7:11115. doi: 10.1038/ncomms11115

Veenma D, Brosens E, de Jong E, van de Ven C, Meeussen C, Cohen-Overbeek T, Tibboel D (2012) Copy number detection in discordant monozygotic twins of congenital diaphragmatic hernia (CDH) and esophageal atresia (EA) cohorts. Eur J Hum Genet 20(3):298–304. doi: 10.1038/ejhg.2011.194

Weber-Lehmann J, Schilling E, Gradl G, Richter DC, Wiehler J, Rolf B (2014) Finding the needle in the haystack: differentiating "identical" twins in paternity testing and forensics by ultra-deep next generation sequencing. Forensic Sci Int Genet 9:42–46. doi: 10.1016/j.fsigen.2013.10.015

Wong C. C. Y., Caspi A, Williams B, Craig IW, Houts R, Ambler A, … Mill J (2010) A longitudinal study of epigenetic variation in twins. Epigenetics 5(6):516–526. doi: 10.4161/epi.5.6.12226